

**Longitudinal and Spatial Analyses Applied to Corn Yield Data from a
Long-Term Rotation Trial - Institute of Statistics Mimeo Series #2559**

Cavell BROWNIE, Larry D. KING, and Tina J. DUBE.

The statistical analysis of yield data from a long term rotation trial presents a number of challenges, as illustrated by a 10 year study on the effects of reduced chemical inputs and different cropping sequences, most of which involved corn (*Zea mays* L.) (King and Buchanan, 1993). Important features of the yield data from this trial include considerable unbalance, possible correlations across time and space, possible heterogeneity in the error variances across years, and the availability of pretreatment measurements of soil properties. We investigated a number of analyses of the rotation trial corn yields, restricting attention to models that could be implemented using standard software, specifically Proc Mixed in SAS (SAS Institute, 1999). These analyses included several types of split plot, repeated measures and spatial analyses, each with and without soil covariates. Diagnostic graphs and AIC values indicated that the model should allow for heterogeneity across years in the error variance. The precision of contrasts in years with high (respectively, low) error variance was underestimated (respectively, overestimated) in analyses that assumed constant variance. Allowing spatial correlations resulted in a small reduction in estimates of precision, and including soil covariates was important in two of nine years. The best models were a repeated measures with banded covariance, as well as autoregressive repeated measures and isotropic spatial models, both modified to allow heterogeneity of the error variance across years. Estimates and *t*-values for specific treatment effects were similar for these models.

Cavell Brownie is Professor and Tina J. Dube a graduate of the Department of Statistics, North Carolina State University, and Larry D. King is Professor Emeritus, Department of Soil Science, North Carolina State University, Raleigh NC 27695, USA.

1. INTRODUCTION

The sustainability of agricultural cropping systems is an increasingly important topic. The only way to compare the cumulative effects of different cropping systems over a number of years is to conduct long-term trials where the treatments involve different crop sequences as well as management practices (e.g., Mueller et al., 2002). Classical rotation trials have long been used to assess sustainability of yield for different crop sequences. A design feature of these classical rotation trials is that each possible ‘start’ is included for each crop rotation to ensure that every phase of each rotation is present every year (e.g., Singh and Jones, 2002; Patterson, 1964). The desire to study a number of management factors as well as different cropping sequences can, however, make it infeasible to include all of the different starts for every rotation (e.g., King and Buchanan, 1993).

Management systems with reduced chemical inputs are of particular interest in assessing sustainability. King and Buchanan (1993) describe a study that was initiated in 1985 to compare two management systems (conventional and reduced chemical inputs) in four cropping sequences (continuous grain sorghum, continuous corn, a 2-year corn rotation and a 4-year, later changed to 3-year, corn rotation). Due to the large number of different cropping systems, not all starts were included for the 2- and 4-year rotations. King and Buchanan (1993) report results of separate analyses of the corn yields from each of the

first seven years of this trial. Analyses of the data, combined across years, have not yet been reported.

Singh and Jones (2002) examined several analyses of yields for two trials that involved 2-year barley rotations in a classical design. Focusing on the repeated measures nature of the data, these authors compared five models for the within-plot error covariance structure, namely PS, the split plot model with constant error variance and constant covariance over years; AR, a first order autoregressive structure; PSH and ARH, generalizations of PS and AR, respectively, to allow for heterogeneity of the error variance across years; and C, a control model with *iid* errors. The models that allowed heterogeneity across years were preferred based on the Akaike Information Criterion (AIC).

To select an appropriate analysis for the corn yield data from the long-term, reduced inputs, rotation trial, we consider several features of the data. As noted by Singh and Jones (2002), the repeated measures nature of the yields is important. Also, separate analyses by year suggested that, as in the barley rotation trials, the error variance differed across years. In addition, precision may be increased by using an analysis that accounts for spatial variation by modeling the correlation between plot errors as a function of the distance between plots (e.g., Brownie, Bowman and Burton, 1993; Cullis and Gleeson, 1991). Precision may also be increased by treating measurements of soil texture and plot elevation, obtained at the start of the trial, as potential covariates. Finally, because not all starts were included for each crop sequence/management combination, the degree of unbalance in yield data for the major crop, corn, is considerable.

There are several ways to account for correlations between errors over time and space, leading to a number of possible models for analyzing the corn yields. In addition to the PS, AR, PSH and ARH models in Singh and Jones (2002), possible covariance structures include banded models for correlations over time, isotropic and anisotropic exponential or power covariance models for correlations over space, and a 3-dimensional space-time covariance structure, all with and without covariates. Determining the best analysis *a priori* is not possible, so our goal is to compare and evaluate a number of plausible models, restricting attention to those that can be implemented using readily available software, specifically Proc Mixed in SAS (SAS Institute, 1999; Littell *et al.*, 1996). More information about the rotation trial is given in Section 2, and the different analyses are outlined in Section 3. Results and discussion follow in Section 4.

2. THE ROTATION STUDY

The rotation trial was initiated in 1985 in an area of about 6 ha, the experimental layout being a randomized block design with 4 blocks. Visual assessments of soil texture and topography were used to determine blocks. As indicated in Figure 1, resulting blocks were irregular in shape and plots in the same block were not always contiguous. Plot dimensions were 8 x 30 m, and coordinates (in m) along east-west and north-south axes were determined for plot centers. After plots had been established, values for clay, gravel, sand and silt content were obtained for each plot, as well as the elevation at each corner of the plot. Figure 1 shows that clay content was generally lower in blocks 3 and 4 than in

1 and 2. Some of the gaps in Figure 1 occur because only plots containing corn rotations are shown. Other gaps are grassed waterways.

Rotations in the long-term study involved four crop sequences: continuous grain sorghum, continuous corn, a 2-year rotation of corn and double-cropped winter wheat and soybean, and a 4-year corn, winter wheat/soybean, corn, red clover hay rotation (changed in 1989 to a 3-year corn, red clover hay, winter wheat/soybean rotation). Crop sequences were studied under two management systems (current management practices, or CMP, and reduced chemical inputs, or RCI). In addition, the CMP rotations involved a third factor, nitrogen (N) fertilizer, at 4 rates. In the RCI systems, N was supplied by a clover green manure crop, or by red clover in the rotation. Two entry points or starts were included for the 2-year corn sequences (though only at the high rate of N in the CMP system). Three starts were included for the 4-year/3-year RCI rotations. By including 'dummy' treatments, provision was made at the beginning of the study to allow for introduction of new treatments as the study progressed. A detailed description of the study can be found in King and Buchanan (1993).

Cropping sequence/management/start combinations are referred to as rotations for convenience. Some rotations were modified during the course of the experiment but after 1989 there were 27 different rotations that involved corn. Corn yields are available for the years 1987-1992, and 1994-1996; the yields in 1993 were poor due to low rainfall, and were not recorded. The number of rotations that provided corn yields in a given year varied between 12 and 25.

3. STATISTICAL ANALYSES

3.1 MODELS AND COVARIANCE STRUCTURES.

To describe the different analyses, we first ignore the soil covariates, and then briefly indicate how to include the covariate information. Ignoring covariates, a linear model for the corn yields is

$$Y_{ijk} = \mu + \beta_i + (RY)_{jk} + \delta_{ij} + \varepsilon_{ijk} \quad (1)$$

where Y_{ijk} is the yield in year k for rotation j in block i , $k = 1, \dots, 6, 8, 9, 10$,

$j = 1, \dots, 27$, $i = 1, \dots, 4$,

β_i is a random effect for the i^{th} block or rep,

$(RY)_{jk}$ is a fixed effect for rotation j in year k , with $\sum_{jk} (RY)_{jk} = 0$,

δ_{ij} is a random effect for the ij^{th} plot (the plot in block i containing rotation j), and

ε_{ijk} is a random error associated with the ij^{th} plot in year k .

The corn yield data are highly unbalanced because of the different crop sequences, with corn planted in 155 of the possible $9 \times 27 = 243$ rotation-year combinations. (There are also several missing yields.) Thus instead of including main and interaction effects for rotation and year, we fit an effect for each observed rotation-year combination, represented as $(RY)_{jk}$ in the linear model (1). Also, the different rotations represent combinations of levels of three factors: crop sequence (with possibly multiple starts), management and fertilization, but only a subset of the possible combinations is used. It is

therefore easier to treat rotations as a single factor and use contrasts for tests relating to the treatment structure.

The models used in analyzing the data are defined by different assumptions about the random effects in the linear model (1). These models, and the commonly used descriptions of the corresponding analyses, are outlined below. Where appropriate, a generalization that allows heterogeneity of the error variance across years is included. Note that SPL, AR and ARH are analogous, respectively, to PS, AR and ARH in Singh and Jones (2002), though SPLH is not the same parameterization as PSH. The SAS code to carry out each analysis is presented in the Appendix.

Unless otherwise noted, block effects β_i are assumed to be *iid* with mean 0 and variance σ_β^2 , and to be independent of all other random effects.

Split plot.

SPL corresponds to a split plot analysis with year as a subplot factor. In the SPL model, the random effects δ_{ij} and ε_{ijk} are assumed to be mutually independent, each with mean 0, and with variances σ_δ^2 and σ_ε^2 , respectively.

SPLH is a modification of the split plot model to allow year specific error variances; δ_{ij} and ε_{ijk} are as for PS, *except* that $Var(\varepsilon_{ijk}) = \sigma_{\varepsilon_k}^2$, $k = 1, \dots, 6, 8, 9, 10$.

Repeated measures with autoregressive errors.

Analysis of repeated measures data using the split plot approach can lead to invalid tests because the split plot model implies that all pairs of measurements on the same plot are equicorrelated, whereas the correlation between two measurements taken over time on the same plot tends to decrease as the time between measurements increases. To reflect

this latter correlation structure, within-plot errors are modeled as a first order autoregressive, or AR(1), process (see also Singh and Jones, 2002).

AR represents a repeated measures analysis with an AR(1) structure for errors in the same plot across time. Plot effects δ_{ij} are as for SPL, and $Cov(\varepsilon_{ijk}, \varepsilon_{ijl}) = \sigma_{\varepsilon}^2 \rho^{|k-l|}$;
 $Cov(\varepsilon_{ijk}, \varepsilon_{i'j'l}) = 0$ if $i'j' \neq ij$.

ARH is a modification of AR that allows year-specific error variances, with

$$Cov(\varepsilon_{ijk}, \varepsilon_{ijl}) = \sigma_{\varepsilon_k} \sigma_{\varepsilon_l} \rho^{|k-l|} .$$

Repeated measures with unstructured error covariance.

UN corresponds to another type of ‘repeated measures’ analysis in which the within-plot error covariance is ‘unstructured’ or as general as possible (e.g., Gumpertz and Brownie, 1993). Thus UN allows heterogeneity of the error variance across years but the required number of covariance parameters is large. Plot effects are subsumed in the error covariance, thus $\delta_{ij} \equiv 0$, and $Cov(\varepsilon_{ijk}, \varepsilon_{ijl}) = \sigma_{\varepsilon_{kl}}$.

Repeated measures with banded error covariance.

UN m , $m = 2, 3, 4, \dots, 8$, is a reduced parameter modification of UN in which the error covariance is assumed to be 0 for measurements in the same plot that are separated by at least m years. $Cov(\varepsilon_{ijk}, \varepsilon_{ijl}) = \sigma_{\varepsilon_{kl}}$ if $|k-l| < m$, $Cov(\varepsilon_{ijk}, \varepsilon_{ijl}) = 0$ otherwise.

Spatial analysis – isotropic covariance.

The effects on yield of small-scale variation in soil can be modeled by assuming that the correlation between effects for two plots is a decreasing function of the distance between the plots. If the correlation depends on distance only, and not on direction, an isotropic

model holds, whereas an anisotropic model is needed if the correlation is not the same in all directions.

SPI assumes an isotropic exponential covariance model for the plot effects δ_{ij} . The errors ε_{ijk} are *iid* representing a nugget effect with constant variance across years. Thus

$Cov(\delta_{ij}, \delta_{lm}) = \sigma_{\delta}^2 e^{-d\theta}$, where d is the distance between plots ij and lm . An

equivalent parameterization that is implemented in SAS with the TYPE = SP(POW)

option (see the Appendix) is obtained by defining $\rho = e^{-\theta}$.

SPIH is a modification of SPI that allows the variance of the errors ε_{ijk} to differ across years, similar to allowing a nugget effect of different magnitude in each year.

Spatial analysis – anisotropic covariance.

SPA assumes that plot effects have the anisotropic covariance structure given in equation

(3.11) of Zimmerman and Harville (1991); that is, $Cov(\delta_{ij}, \delta_{lm}) = \sigma_{\delta}^2 \rho_1^{d_1} \rho_2^{d_2}$, where

d_1 and d_2 are the distances between plots ij and lm in the east-west and north-south

directions, respectively. As noted in the Appendix, this is implemented using the TYPE = SP(POWA) option.

SPAH is a modification of SPA that allows the variance of the errors ε_{ijk} to differ across years.

Spatio-temporal error covariance.

SP3D models the correlations across space and across time using a 3-dimensional

anisotropic covariance structure for the errors. This covariance is related to the

parsimonious variogram model proposed by Ersboll (1997) and also to equation (3.11) in

Zimmerman and Harville (1991). The plot effects are subsumed in the model for the

errors which allows for 2-dimensional spatial variation with time as the third dimension.

Thus, $\delta_{ij} \equiv 0$, and $Cov(\varepsilon_{ijk}, \varepsilon_{i'j'k'}) = \sigma_\varepsilon^2 \rho_1^{d_1} \rho_2^{d_2} \rho_3^{d_3}$, where d_1 and d_2 are the distances between plots ij and $i'j'$ in the east-west and north-south directions, respectively, and $d_3 = |k - k'|$, the number of years between the observations.

3.2 INCLUDING SOIL COVARIATES.

Each of the above analyses can be modified by including one or more of the baseline soil measurements or plot elevation as covariates. These soil properties are true covariates in that they were measured at the start of the study and represent properties (e.g., clay content) that are not likely to be modified over time by effects of the rotations.

Depending on the amount of rainfall, however, the relationship between yield and, e.g., clay content, can differ across years. This was explored by obtaining residuals after fitting rotation by year effects and graphing residuals against each soil property separately for each year. Such graphs, and other preliminary analyses, suggested a model with clay content and gravel content as covariates, with different slope parameters each year. One way to represent this model is:

$$Y_{ijk} = \mu + \beta_i + (RY)_{jk} + \gamma_{1k}C_{ij} + \gamma_{2k}G_{ij} + \delta_{ij} + \varepsilon_{ijk} \quad (3)$$

where C_{ij} and G_{ij} represent clay and gravel content, respectively, measured at baseline, for plot ij , and γ_{1k} and γ_{2k} represent the slopes in the regression of yield in year k on clay and gravel content, respectively. Assumptions concerning the random effects do not change.

3.3 TESTS FOR ROTATION EFFECTS.

After selecting an appropriate analysis, within- and between-year comparisons among rotations can be tested using Estimate statements. For illustration we present results for a

limited number of interesting comparisons. King and Buchanan (1993) note that in the RCI management system, the first year when the Rotation effect becomes significant is 1992. We therefore compared the mean of the 2- and 4-/3-year rotations under RCI management (R2 and R4) with the RCI continuous corn (R1) in each of the years, and averaged over years. We also tested for a linear trend across years in this contrast. Other contrasts tested were R2, R4 compared to the CMP continuous corn at the 0 and 70 kg ha⁻¹ rates of N (C1 and C2), both within years and averaged over years.

Within-year comparisons can, of course, be tested in separate analyses by year, but across-year comparisons, must be tested in a combined analysis. A model that assumes constant error variance may produce inappropriate standard errors (SEs), especially for within-year contrasts, if the homogeneity assumption fails.

4. RESULTS AND DISCUSSION

AIC values are presented for models without soil covariates in Table 1, and for models with covariates included in Table 2. As the fixed effects are different for these two sets of models, the AIC values are used to compare models within a table but not between tables. Not surprisingly, given the way that blocks were determined, including clay and gravel as covariates with year-specific slopes resulted in an estimate of 0 for σ_{β}^2 for all models. Thus each model in Table 2 was fitted with σ_{β}^2 assumed to be 0, and had one fewer covariance parameter than the corresponding model in Table 1.

Estimates of the error variance $\sigma_{\varepsilon_k}^2$ obtained with the UN model (ignoring covariates) varied from 0.33 in 1987 to 1.02 and 1.12 in 1995 and 1989, respectively, while σ_{ε}^2 was estimated to be 0.60 with SPL (Figure 2). Including covariates reduced the error variance under SPL to 0.53, and resulted in little change under UN for $\hat{\sigma}_{\varepsilon_k}^2$ for 1987-1991.

Reductions under UN were more substantial in the later years, particularly 1995, where including covariates reduced $\hat{\sigma}_{\varepsilon_k}^2$ from 1.02 to 0.53. Whether or not covariates were included, the pattern of variation across years in $\hat{\sigma}_{\varepsilon_k}^2$ under ARH, the banded models, SPIH and SPAH, was similar to that for UN.

AIC values for SPLH, ARH, SPIH, and SPAH in Tables 1 and 2, support the visual impression in Figure 2 that there is heterogeneity in the error variance across years. Each of these models is preferred to (has a lower AIC value than) the more restrictive counterpart that assumes a constant error variance across years. The banded UN m models, which allow heterogeneity across years but ignore spatial correlations, have low AIC values for $m = 3, 4, 5$. (Though not in the tables, the AIC values for UN5 are 1286.6 and 1346.9.) Compared to UN3, UN4 and UN5, the UN model appears to be overparameterized, and UN2 is underparameterized.

The correlation structure (ignoring block effects) for yields in the same plot across years is compared in Figure 3 for the SPL, AR and UN models, fitted without covariates.

Results are similar if covariates are included. Estimated correlations are graphed against the lag, or the number of years between yields. The correlation under SPL does not

depend on the lag and is estimated as $\hat{\sigma}_{\delta}^2 / (\hat{\sigma}_{\delta}^2 + \hat{\sigma}_{\varepsilon}^2) = .105 / (.105 + .598) = 0.15$. The

AR correlation depends on the lag d and is obtained as $(\hat{\sigma}_\delta^2 + \hat{\sigma}_\varepsilon^2 \hat{\rho}^d) / (\hat{\sigma}_\delta^2 + \hat{\sigma}_\varepsilon^2)$, where $\hat{\sigma}_\delta^2 = 0.066$, $\hat{\rho} = 0.297$ and $\hat{\sigma}_\varepsilon^2 = 0.642$. Correlations under the UN model (which places no restrictions on the structure) appear to decrease in absolute magnitude as the lag increases but are highly variable at small lags. The pattern of the estimated correlations under UN is consistent with the low AIC values for the banded models UN3, UN4, UN5, which force a correlation of 0 at high lags and allow a general structure at low lags. These banded models are preferred to SPL, AR and UN2 (each of which has a smaller number of parameters than UN3), and are also preferred for reasons of parsimony to the UN model.

To assess the correlations across space, empirical variograms were computed on estimated plot effects and separately for each year on residuals obtained after fitting rotation effects. These variograms were generally similar in shape with an effective range equal to 2 or 3 plot lengths. The sill appeared to depend on year, however, agreeing with the lower AIC values for SPIH and SPAH in comparison to the SPI and SPA models. The AIC values for these four models also support use of the isotropic, rather than anisotropic, spatial structure. Harder to explain are the low AIC values for the 3-dimensional space-time model, SP3D, which does not allow heterogeneity of the error variance across years.

Results for treatment contrasts are given in Table 3 for ARH, UN3, SPIH, SP3D (four of the top five models in terms of AIC), and also for SPL. For brevity, estimates and SEs are reported only for the R2, R4 vs R1 contrast in selected years, and for the linear trend across years. In general, except for SP3D, the analyses produce similar point estimates of effects. With respect to SEs for the within-year contrasts R2, R4 vs R1, results for SP3D

agree more closely with results for SPL than with the models that allow variance heterogeneity (ARH, UN3 and SPIH). Thus, under SPL and SP3D, SEs appear to be underestimated for 1989, and overestimated for 1992, years with, respectively, high and low values of $\hat{\sigma}_{\varepsilon_k}^2$. The same pattern holds for the within-year contrasts R2, R4 vs C1 and R2, R4 vs C2.

Results are more similar across analyses for effects averaged over years than for within-year effects (though SP3D consistently produces the smallest SEs). For the linear trend across years in the R2, R4 vs R1 contrast, SEs differ across analyses and SPIH has the smallest SE. Differences in the fitted covariance structure across time and space may have a greater impact for the linear effect, compared to the average effect, because early years (including 1989) and late years are weighted more heavily than middle years.

Fitting covariates noticeably reduces error for contrasts in 1992 and 1995 for ARH, UN3 and SPIH but not for SP3D and SPL. The effect of fitting covariates is smaller for SEs of effects averaged over years. Compared to the split plot and repeated measures analyses, p-values for the covariates clay and gravel are larger (though significant at .05) under the spatial analyses, indicating some overlap between the two methods (use of covariates versus spatial correlations) of accounting for within-block variation.

For the UN3 model with covariates included, averaging over years, yields for the RCI 2- and 4-/3-year rotations (R2, R4) were higher than for RCI continuous corn (by 0.8 ± 0.23 kg ha⁻¹), higher than for CMP continuous corn with 0 N (by 1.9 ± 0.26 kg ha⁻¹), but lower than for CMP continuous corn with 70 kg ha⁻¹ N (by 0.9 ± 0.26 kg ha⁻¹). In the RCI system, the additional yield with the 2- and 4-/3-year rotations, compared to continuous corn, increased on average by $.12 \pm 0.05$ kg ha⁻¹ per year over the 10 year

period. Similar conclusions are reached with each of the analyses, though the linear effect is not significant under SP3D.

The agreement between SEs for SPL and SP3D suggests that even though the AIC value is low for SP3D, given the evidence of heterogeneity, this model is not appropriate.

Assuming heterogeneity is present, another reason that p-values for within-year contrasts will be inaccurate under SP3D, is that error degrees of freedom (using the Kenward-Roger method, Kenward and Roger, 1997), are too large, because the effect of estimating an error variance for each year is not taken into account.

The models investigated are by no means the only plausible error covariance structures for this data set. For example, an alternate spatial model assumes that plot effects are *iid* and the errors ε_{ijk} are spatially correlated within each year. With this parameterization, including heterogeneity in the spatial covariance across years gave AIC values larger than for SPIH and SPAH. Because our goal was to focus on analyses that could be implemented using standard software, more complex spatial models (*e.g.*, the semiparametric models of Durban, Hackett, McNichol *et al.*, 2003) or space-time models were not considered. Such models would have to take into account the unbalance, strong year effects, and heterogeneous error variance.

Our results reinforce the findings of Singh and Jones (2002) that accounting for heterogeneous errors across years is important for long-term trials. Of the covariance models examined for analysis of the yield data across years, the most appropriate are ARH, the banded UN3 and UN4, and also SPIH. Overall, fitting soil covariates did increase precision, but the effect was minor in most years.

APPENDIX

PROC MIXED CODE.

The code needed to implement the analyses in SAS Proc Mixed is outlined. Fixed effects are listed in the Model statement and random effects in the Random and Repeated statements. Plot effects, δ_{ij} , are indexed by all combinations of Block and Rotation (i.e., by BLOCK*ROTATION). The structure for the δ_{ij} is specified in the Random statement, while that for the errors ε_{ijk} is in the Repeated statement. If no specific structure is given via the TYPE or GROUP options, levels of a random effect are treated as *iid*.

```
PROC SORT; BY BLOCK ROTATION YEAR;
***** SPL or split plot *****;
PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK BLOCK*ROTATION;
LSMEANS ROTATION*YEAR;
***** SPLH, split plot with year-specific error variances *****;
***** obtained via the GROUP option in Repeated *****;
PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK BLOCK*ROTATION;
REPEATED BLOCK*ROTATN*YEAR / GROUP = YEAR SUBJECT = BLOCK*ROTATN;

***** AR or AR(1) errors *****;
PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK BLOCK*ROTATION;
REPEATED YEAR / TYPE = AR(1) SUBJECT = BLOCK*ROTATION;
```

***** ARH or AR with heterogeneous error variance *****;
*** Replace TYPE = AR(1) with TYPE = ARH(1) in Repeated statement for AR ***;

***** UN or unstructured error covariance *****;
PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK;
REPEATED YEAR / TYPE = UN SUBJECT = BLOCK*ROTATION;

***** UN2 or banded error covariance *****;
*** Replace TYPE = UN with TYPE = UN(2) in the Repeated statement for UN *****;
***** Similarly for UN3, UN4, etc. *****;

***** SPI or isotropic spatial covariance *****;
*** WEST and SOUTH represent the plot coordinates measured in the east-west
*** and north-south directions, respectively. *****;

PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK;
RANDOM BLOCK*ROTATION / TYPE = SP(POW) (WEST SOUTH)
SUBJECT = INTERCEPT;

*** For SPA, change POW to POWA in the second Random statement for SPI ***;

***** SPIH - isotropic spatial with heterogeneous nugget *****;
PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK;
RANDOM BLOCK*ROTATION / TYPE = SP(POW) (WEST SOUTH)
SUBJECT = INTERCEPT;

REPEATED BLOCK*ROTATION*YEAR / GROUP = YEAR ;

For SPAH, change POW to POWA in the second Random statement for SPIH;

***** SP3D – anisotropic space-time *****;
PROC MIXED; CLASS BLOCK ROTATN YEAR;
MODEL YIELD = ROTATN*YEAR / DDFM = KENWARD;
RANDOM BLOCK;

```
REPEATED / TYPE = SP(POWA) (WEST SOUTH YEAR) SUBJECT = INTERCEPT;  
RUN;
```

Notes:

(i) To include soil covariates, the Model statement is modified. For example, to allow for regression on CLAY = clay content and GRAV = gravel content, with year-specific coefficients, the Model statement is:

```
MODEL YIELD = ROTATN*YEAR CLAY CLAY*YEAR GRAV GRAV*YEAR  
/ DDFM = KENWARD;
```

(ii) When covariates were included, σ_{β}^2 was estimated to be 0 for all models, so the covariance analyses were run with the Model statement in (i) above and with BLOCK deleted from the Random statement.

(iii) The TYPE = AR(1) and ARH(1) options in Proc Mixed assume equally spaced measurement times. This presents a problem with the corn yields because there is a 2-year gap in the data between 1992 and 1994. Including 1993 yields as missing values in the data set produces the correct estimates with AR(1). Results with ARH(1) also seem correct though the output contains a meaningless estimate of the error variance for 1993.

(iv) To obtain estimates of mean yields for Rotations in different years for a given analysis, an Lsmeans statement is added as shown in the code for SPL. Estimate or Contrast statements can be used to test for within- and across-year contrasts on the Rotations. There are 155 Rotation*Year combinations, listed first by rotation and then by year, so care must be taken to specify the coefficients correctly.

(v) Some models took more than 30 minutes of CPU time on a PC.

REFERENCES

- Brownie, C., Bowman, D.T., and Burton, J. (1993), "Estimating Spatial Variation in Analysis of Data from Yield Trials: A comparison of methods," *Agronomy Journal* 85, 1244-1253.
- Brownie, C. and Gumpertz, M. (1997), "Validity of Spatial Analyses for Large Field Trials," *Journal of Agricultural Biological and Environmental Statistics* 2:1-23.
- Cullis, B. R. and Gleeson, A. C. (1991), "Spatial Analysis of Field Experiments – an Extension to Two Dimensions," *Biometrics* 47, 1449-1460.
- Ersboll, A. K. (1997), "A Comparison of Two Spatio-temporal Semivariograms with Use in Agriculture," in *Modelling Longitudinal and Spatially Correlated Data: Methods, Applications and Future Directions*, T. G. Gregoire *et al.* eds. Pp299-308.
- Durban, M. Hackett, C. A., McNicol, J. W., Newton, A. C., Thomas, W. T. B., and Currie, I. A. (2003), "The Practical use of semiparametric models in Field Trials," *Journal of Agricultural, Biological and Environmental Statistics* 8, 48-66.
- Kenward, M. G. , and Roger, J. H. (1997), "Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood," *Biometrics* 53, 983-997.
- King, L. D., and Buchanan, M. (1993), "Reduced Chemical Input Cropping Systems in the Southeastern United States. I. Effect of Rotations, Green Manure Crops and Nitrogen Fertilizer on Crop Yields," *American Journal of Alternative Agriculture*, 8, 58-77.
- Littell, R. C., Milliken, G. A., Stroup, W. W., and Wolfinger, R. D. (1996), *SAS System for Mixed Models*, New York: SAS Institute, Inc.
- Mueller, J.P., Barbercheck, M.E., Bell, M. Brownie, C., Creamer, N.G., *et al.* (2002), "Development and Implementation of Long-term Agricultural Systems Studies: Challenges and Opportunities," *Horttechnology* 12:362-368.
- Patterson, H. D. (1964), " The Theory of Cyclic Rotation Experiments" (with discussion), *Journal of the Royal Statistical Society*, Series B, 26, 1-45.
- SAS Institute, Inc. (1999), SAS version 8, Online Help, Cary, NC: SAS Institute.

Singh, M. and Jones, M. J. (2002), "Modeling Yield sustainability for Different Rotations in Long-Term Barley Trials," *Journal of Agricultural, Biological and Environmental Statistics* 7, 525-536.

Zimmerman, D. L. and Harville, D. A. (1991), "A Random Field Approach to the Analysis of Field-Plot Experiments and Other Spatial Experiments," *Biometrics* 47, 223-239.

Table 1. AIC values for models with split plot, repeated measures and spatial covariance structures, and without fitting covariates.

Model	# of parameters [†]	AIC	Model	# of parameters [†]	AIC
SPL	3	1325.1	SPLH	11	1301.5
AR	4	1317.9	ARH	12	1289.7
UN2	18	1303.7	UN3	25	1284.3
UN4	31	1286.4	UN	46	1300.3
SPI [#]	4	1319.5	SPIH [#]	12	1294.1
SPA [#]	5	1320.9	SPAH [#]	13	1294.8
SP3D	5	1285.8			

[†] the number of covariance parameters in the fitted model, including σ_{β}^2 , even in cases where this parameter is estimated to be 0 .

[#] σ_{β}^2 is estimated to be 0 .

Table 2. AIC values for models with split plot, repeated measures and spatial covariance structures, with clay and gravel content as covariates.

Model	# of parameters [†]	AIC	Model	# of parameters [†]	AIC
SPL	2	1373.2	SPLH	10	1351.4
AR	3	1364.3	ARH	11	1337.1
UN2	17	1351.8	UN3	24	1341.5
UN4	30	1343.6	UN	45	1362.9
SPI	3	1369.2	SPIH	11	1349.2
SPA	4	1370.5	SPAH	12	1350.4
SP3D	4	1343.8			

[†] σ_{β}^2 was estimated to be 0 for all analyses. Therefore all models were refitted assuming $\sigma_{\beta}^2 = 0$, and the number of covariance parameters is one less than that for the corresponding model in Table 1.

Table 3. Estimates (and SEs) of RCI rotation effects on mean yield (kg ha⁻¹) under five methods of analysis. The difference between the RCI 2- and 4-/3-year rotations (R2 and R4) and RCI continuous corn (R1) is reported for selected years, as well as the average increase per year in this comparison.

Without Covariates					
Contrast	SPL	ARH	UN3	SPIH	SP3D
R2, R4 vs R1 1989	0.58 (0.36)	0.58 (0.47)	0.58 (0.46)	0.53 (0.45)	0.43 (0.31)
R2, R4 vs R1 1992	0.80 (0.51)	0.80 (0.40)	0.80 (0.44)	0.84* (0.40)	0.75 (0.48)
R2, R4 vs R1 1995	1.57** (0.51)	1.57* (0.64)	1.57* (0.60)	1.55* (0.61)	1.42** (0.44)
R2, R4 vs R1 1996	1.19* (0.51)	1.19* (0.53)	1.19* (0.50)	1.24* (0.51)	0.85 (0.49)
R2, R4 vs R1 Linear	0.12* (0.050)	0.12 (0.060)	0.11* (0.053)	0.12* (0.049)	0.09 (0.055)
With Covariates					
Contrast	SPL	ARH	UN3	SPIH	SP3D
R2, R4 vs R1 1989	0.59 (0.35)	0.57 (0.48)	0.57 (0.47)	0.57 (0.46)	0.35 (0.31)
R2, R4 vs R1 1992	1.05* (0.49)	1.06** (0.36)	1.08** (0.38)	1.06** (0.37)	0.96* (0.47)
R2, R4 vs R1 1995	1.40** (0.49)	1.40** (0.48)	1.41** (0.45)	1.38** (0.45)	1.39** (0.43)
R2, R4 vs R1 1996	1.17* (0.49)	1.19* (0.53)	1.17* (0.49)	1.18* (0.49)	0.90 (0.48)
R2, R4 vs R1 Linear	0.12* (0.047)	0.12* (0.055)	0.12* (0.050)	0.12* (0.045)	0.11 (0.054)

* indicates P < .05, ** indicates P < .01.

FIGURE CAPTIONS

Figure 1. Field layout for the reduced chemical inputs rotation trial showing the four blocks, with plots indicated by block number, and by percent clay content of the soil. Only plots with corn rotations are shown.

Figure 2. Estimated error variance for each year with the UN model, ignoring covariates (○) and with covariates fitted (+). For comparison, the estimated (constant) error variance using the SPL model, ignoring covariates (Δ) and with covariates fitted (◇), is also shown.

Figure 3. Correlations between errors across time in the same plot in relation to the Lag, or number of years between measurements, estimated by fitting an unstructured error covariance (○), by fitting an AR(1) covariance (●), and the (constant) correlation with the SPL model (Δ). All estimates are from analyses ignoring soil covariates.



