MODELLING WITHIN-PLANT SPATIAL DEPENDENCIES OF COTTON YIELD

E. B. Moser

R. E. Macchiavelli

D. J. Boquet

MODELLING WITHIN-PLANT SPATIAL DEPENDENCIES OF COTTON YIELD

E. B. Moser[1], R. E. Macchiavelli[2], and D. J. Boquet[3]

[1]Dept. Experimental Statistics and Louisiana Ag. Exper. Sta., Louisiana State University and LSU Ag. Center, Baton Rouge, LA 70803-5606, [2]Dept. Experimental Statistics, LSU, and [3]Louisiana Ag. Exper. Sta., Northeast Research Sta., P.O. Box 438, St. Joseph, LA 71366

Abstract

In field experiments during 1987-1990, cotton plants were grown under 8 different levels of nitrogen application to assess the impact of nitrogen fertilization on the fruiting and yield distribution of cotton within the plant (Boquet et al. 1993). Individual boll weights and average seedcotton yield were determined at each fruiting site for each main-stem node along the plant. Various models of dependence and independence are possible to explain and account for the dependencies of the yields among the sites and nodes of the plant. Here we investigate models of total yield per node and yield per node adjusted for the number of sites using several models for the spatial dependence among the nodes. Typical univariate models would either assume a simple homogeneous error structure or a compound symmetry error structure among the nodes, leading to the split-plot-type models. A multivariate unstructured approach ignores obvious spatial dependencies among the nodes. Spatial models and ante-dependence models permit a parsimonious summary of the error structure and are compared with the compound symmetry and multivariate models.

Keywords: Geostatistics, Spatial Models, Variogram, Repeated Measures

1. Introduction

The capability to model and test hypotheses concerning repeated measures data has advanced rapidly over the past 10 years due primarily to the development of fast computing hardware with large storage availability and to software development to handle the non-linear optimizations necessary for the estimation of model parameters. Previously, analysis of continuous repeated measures data made certain assumptions about the correlation or covariance structure among the set of repeated measurements, such as independence among measurements, equal correlations among the set of repeated measurements, which is also known as compound symmetry, or a completely unstructured dependence, which is the general multivariate analysis of repeated measures (see e.g., Crowder and Hand 1990, Moser et al. 1990, Lindsey 1993). The independence model might be appropriate if the measurements are made sufficiently far apart in space or time such that a dependence among them is very minor. This, however, is usually counter to the reasons for designing an experiment or study using repeated measurements. The compound symmetry model specifies a dependency structure but also assumes that the correlation among measurements made close in time or space have the same correlation as measurements

made far apart in time or space. This model might be appropriate for split-plot designed experiments where randomization leads to equal correlations among the treatment levels within main-plot units. This also includes some repeated measures experiments where time or space is not the repeated measures factor. However, most repeated measures experiments have restricted randomization since time or space cannot be reordered, and the dependence is related to the separation of the measurements in time or space. Lastly, the unstructured case handles any dependencies that might be present, including independence and compound symmetry as special cases, but at the same time does not exploit any structure that might be present in the ordering and spacing of the measurements. This can be important for several reasons. First, the number of independent observations or subjects in the experiment may be small relative to the number of parameters to be estimated in the unstructured case leading to low power for testing hypotheses about treatment effects. Second, the number of repeated measures (points in time or space) may be larger than the number of independent subjects leading to a singular estimated variance-covariance matrix and no tests of hypotheses about the treatment effects.

This paper will explore alternatives to the traditional repeated measures models as presented above and will look at criteria for choosing among them. These alternative models specify structures for the covariance matrix, and therefore, impose structure on the correlations among the set of repeated measurements, or they specify structure for the concentration matrix (inverse of the covariance matrix), which imposes structure on the partial correlations among the set of repeated measurements. These models typically use few parameters for the specification of the dependence and heterogeneity structure, especially compared with the unstructured multivariate model, and so lose fewer degrees of freedom while providing a parsimonious model for the association among the set of repeated measurements.

2. The Data Set

Field experiments were conducted during 1987 to 1990 at the Northeast Research Station near St. Joseph, LA, a unit of the Louisiana Agricultural Experiment Station, Louisiana Agricultural Center, to study the effects of nitrogen fertilization (NRATE) on fruiting and yield distribution of cotton within the plant (Boquet et al. 1993). Many studies have focused on the effects of fertilizer nitrogen rates on total cotton yields or average boll weights per plant. This study was undertaken to learn more about how in-field nitrogen fertilization affects the distribution of cotton within a plant.

The experimental design was a randomized complete block with four blocks containing eight NRATE levels. Individual plots consisted of eight, 19-m long rows with a 1 m inter-row spacing. The treatments were preplant NRATE levels of 0, 28, 56, 84, 112, 140, and 168 kg ha^{-1}, NRATE=1,2,...,7, and a preplant plus sidedress application of 56 + 56 kg nitrogen ha^{-1}, NRATE=9. All fertilizer treatments were maintained on the same plots for the four years of the study. Twenty randomly selected plants within each plot were identified during mid-July and during harvest time these plants were hand-harvested and the location of the boll on the plant was recorded. Boll location was identified according to the

sympodial branch number, referred to as NODE in this paper, and to the fruiting site within the branch (FS) (see Munroe and Farbrother 1969). Average boll weight was determined by dividing the total seedcotton weight by the actual number of bolls harvested at a given fruiting site on a given plot. For the purposes of this paper, an average was computed for each NODE and the effect of fruiting site within NODE was ignored. In the original analysis by Boquet et al. (1993) a strip-plot model using compound symmetry structures was used to consider both FS and NODE, but where NODE had been composited to 1 of 3 plant horizons. Only the 1990 data are used here.

3. Covariance and Concentration Matrix Structures

Statistical analysis and modelling were accomplished using SAS software including PROC MIXED in SAS/STAT (SAS Institute 1992) and PROC IML in SAS/IML (SAS Institute 1989). PROC MIXED can fit a variety of covariance structures using maximum likelihood (ML) or restricted maximum likelihood (REML) and was used to fit all but the ante-dependence structures. We modelled the data as

$$\underset{\sim}{Y}_{ij} = \underset{\sim}{\mu} + \underset{\sim}{\tau}_i + \underset{\sim}{\beta}_j + \underset{\sim}{\varepsilon}_{ij} \tag{1}$$

where $\underset{\sim}{Y}_{ij}$ is a vector of the average boll weights on the 11 NODEs, $\underset{\sim}{\mu}$ is the overall mean vector, $\underset{\sim}{\tau}_i$ is the vector of effects of the ith level of NRATE, $i = 1...8$, $\underset{\sim}{\beta}_j$ is the vector of the jth block effect, $j = 1...4$, and $\underset{\sim}{\varepsilon}_{ij}$ is the vector of residuals from the ith NRATE level of the jth block. This gives an 11×11 covariance matrix for the $Var(\underset{\sim}{Y}_{ij})$ and $Var(\underset{\sim}{\varepsilon}_{ij})$. Notice that in this model we are using a narrow inference space for the blocks so they appear like a fixed effect in the analysis. We will be assuming that $\underset{\sim}{Y}_{ij} \sim N_{11}(\underset{\sim}{\mu} + \underset{\sim}{\tau}_i + \underset{\sim}{\beta}_j, \Sigma)$, which implies $\underset{\sim}{\varepsilon}_{ij} \sim N_{11}(\underset{\sim}{0}, \Sigma)$. We could also write the above multivariate model as a univariate model

$$Y_{ijk} = \mu + \tau_i + \beta_j + \delta_{ij} + \nu_k + \varepsilon_{ijk} \tag{2}$$

where the δ_{ij} are the subject or plant residuals, ν_k is the effect of the kth node, $k = 8...18$, and now ε_{ijk} are the residuals from each node within plants. PROC MIXED uses the model (2) formulation and fits the δ_{ij} using the "SUBJECT=" option of the REPEATED statement. The REPEATED statement is also used to impose a covariance structure on the ε_{ijk} using the "TYPE=" option (SAS Institute 1992).

Let I be an 11×11 identity matrix and let J be an 11×11 matrix with all elements 1. Several models for the covariance, $Var(\underset{\sim}{\varepsilon}_{ij}) = \Sigma$, or concentration matrix, Σ^{-1}, structure were tried including simple,

$$\Sigma = \sigma^2 I, \tag{3}$$

compound symmetry,

$$\Sigma = \sigma^2 I + \theta J, \tag{4}$$

and completely unstructured,

$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,11} \\ & \sigma_{2,2} & \cdots & \sigma_{2,11} \\ & & \cdots & \\ & & & \sigma_{11,11} \end{bmatrix}, \tag{5}$$

which are the structures discussed earlier. We also fit the first-order autoregressive structure,

$$\Sigma = \frac{\eta}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{10} \\ & 1 & \rho & \cdots & \rho^9 \\ & & & \cdots & \\ & & & & 1 \end{bmatrix}, \tag{6}$$

where ρ is the first-order autocorrelation parameter. Since this is a model of conditional independence, the concentration matrix (Σ^{-1}) has a very insightful structure,

$$\Sigma^{-1} = \frac{1}{\eta} \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ & & & \cdots & 0 & 0 \\ & & & & 1+\rho^2 & -\rho \\ & & & & & 1 \end{bmatrix}, \tag{7}$$

which has as the first minor diagonal $-\rho$ and all other minor diagonals are zero. The zeros imply conditional independence between those measurements given the intermediate measurements. The Cholesky decomposition $\Sigma^{-1} = A'A$ has the structure

$$A = \frac{1}{\sqrt{\eta}} \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}, \tag{8}$$

which we will show for the fitted data.

The rth-order ante-dependence structures can also be defined in terms of the Cholesky decomposition $\Sigma^{-1} = A'A$. Define A to have elements a_{ij} such that

$$a_{ij} = \begin{cases} 1 / \sqrt{\delta_i} & j = i \\ -\gamma_{i-j, i} & i - r_i \le j < i \\ 0 & \text{elsewhere} \end{cases} \tag{9}$$

for $i, j = 1 \dots 11$. For a constant-order ante-dependence model of order s, $r_i = \min(i - 1, s)$, while a variable-order ante-dependence model (Macchiavelli and Arnold 1994) is only restricted such that $0 \le r_i \le i - 1$. The concentration matrix for the ante-dependence model looks very similar to that of the autoregressive structure except that the parameters along the main diagonal and along the minor diagonals are not constrained to be the same. Here we assumed that the dependence was such that lower nodes were affecting boll weights higher on the plant but the model will fit exactly the same if we assume the dependence is from top to bottom. The ante-dependence model does require, however, that the ordering of the nodes from top to bottom or bottom to top be preserved. Although PROC MIXED does not fit the general ante-dependence structures, they can be fit using conditional regression methods (see Macchiavelli 1992, Macchiavelli and Arnold 1994) and then the estimated covariance parameters used by PROC MIXED for testing.

Both the autoregressive model and the ante-dependence model permit a smooth decay of the dependence among the repeated measurements with increasing distance or time. The ante-dependence model, however, permits different conditional variances and partial associations at each time period. This permits modelling of unequally spaced time periods or locations, while the autoregressive model will require equal spacing due to the homogeneous parameter structure. The ante-dependence model is also nice in that it gives us a reasonable way of incorporating time-varying covariates in the covariance structure (Macchiavelli and Moser 1994).

Spatial models from geostatistics were also fit as these too provide explanations of the covariance structure as a function of the distance or time separating repeated measurements. Let d_{ij} be the distance between two nodes i and j where we define $d_{ij} \equiv |i - j|$, and let σ_{ij} be the i, jth element of Σ. The geostatistical or spatial covariance structures fit included the linear structure

$$\sigma_{ij} = \sigma^2 (1 - \alpha d_{ij}) \tag{10}$$

for $\alpha d_{ij} \le 2$ and zero otherwise, the linear log

$$\sigma_{ij} = \sigma^2 (1 - \alpha \log(d_{ij})) \tag{11}$$

for $\alpha \log(d_{ij}) \leq 2$ and zero otherwise, the exponential

$$\sigma_{ij} = \sigma^2 \left(\exp(-d_{ij}/\alpha) \right), \tag{12}$$

the Gaussian

$$\sigma_{ij} = \sigma^2 \left(\exp(-d_{ij}^2/\alpha^2) \right), \tag{13}$$

the spherical

$$\sigma_{ij} = \sigma^2 \left(1 - (3d_{ij}) / (2\alpha) - d_{ij}^3 / (2\alpha^3) \right) \tag{14}$$

for $d_{ij} \leq \alpha$ and zero otherwise, and the power law model

$$\sigma_{ij} = \sigma^2 \alpha^{d_{ij}}. \tag{15}$$

For the exponential, Gaussian, and spherical models, the parameter σ^2 is called the sill and α is called the range using geostatistical nomenclature. We can also add a nugget parameter θ^2 to each spatial structure using the LOCAL option of the REPEATED statement in PROC MIXED. Although PROC MIXED of SAS/STAT can fit each of these spatial models, macros written in SAS/IML were also used to provide some diagnostic information for these spatial models.

4. Results

The observed correlation matrix of residuals from the multivariate model is presented in Table 1 (upper triangle) and suggests some interesting relationships among the nodal boll weights. It appears that there is generally a positive association among the residuals of adjacent nodes while residuals separated a great distance have little linear association. The partial correlations given intermediate nodes (Table 1, lower triangle) also suggest that much of the dependency among nodes can be explained by their relationship with nearby nodes. However, if we look at the rows corresponding to nodes 14 through 17 we see that the partial correlations indicate that dependencies for these nodes can extend to the lowest nodes. This suggests a variable order ante-dependence model.

The Cholesky decomposition of the concentration matrix (Table 2) indicates that a conditional independence model accounting for first-order relationships is probably needed and that the conditional variances are not highly unequal. Note that many off-diagonal elements in the lower-triangle are near zero suggesting that some type of conditional independence model could be appropriate. The table of partial correlations given intermediate values (Table 1, lower triangle) is a better device for helping with the model selection as the concentration matrix gives partial covariances given all other nodes. The observed concentration matrix, however, could be useful in examining model fit.

To help choose among the covariance structures, penalized likelihood criteria were used. The Akaike's AIC and Schwarz's BIC criteria were shown by Macchiavelli (1992) to be useful for selecting the order of ante-dependence and have been used routinely for model selection in general. We can write these criteria in the form

$$\text{IC}(k) = \log L_k - \frac{d_k}{2} c(n) \tag{16}$$

where k indexes the model under consideration, d_k is the number of covariance parameters in the kth model, $\log L_k$ is the log-likelihood evaluated at the maximum likelihood estimates, and $c(n) = 2$ for AIC and $c(n) = \log n$ for BIC. For this data set $n = 32$, which is the number of independent observations in the data set (the number of plots or "plants"). The values for AIC and BIC for the various models fit are given in Table 3. The variable-order ante-dependence model appearing in the table is the BIC-optimal one among all variable-order ante-dependence models. Note that the values of BIC reported by PROC MIXED differ from those reported here as the formula used therein takes n to be the total number of observations used in the analysis rather than the number of independent observations. Also note that the ordering of the models based upon the information criteria does not need to be the same as the penalties are quite different. To help discriminate among the various spatial models the empirical semi-variogram was constructed using SAS/IML routines (Figure 1). The figure confirms that any of the spatial models that include a nugget effect could be appropriate, although one might wish to force some asymptote (the sill) on these data. We also examined plots of residuals as a function of the fixed effects to look for any outliers or other troublesome features. We did encounter some numerical problems with PROC MIXED in the estimation and stability of the parameters for some of the models, but we believe that the likelihood values reported are reasonably accurate.

The linear with nugget spatial structure was used to compute least-squares means and to test model reduction hypotheses. This model was selected based upon the BIC criterion but other structures could also have been used (see the discussion below). The overall tests for the fixed effects hypotheses indicate that an interaction between NRATE and NODE is occurring although the F value is not very large (Table 4). To further investigate these effects, various fixed-effects models were constructed for fitting linear and quadratic terms for NRATE, NODE, and their interaction. Quasi-Type II hypotheses were constructed by using the HTYPE=1 option on the PROC MIXED MODEL statement, appropriate ordering of the model effects, and then making several runs under different orderings. These results are given in Table 4 under the Type I F column. Although they are labelled under the TYPE I F column, the NRATE terms were fitted following NODE, and the NODE terms were fit after the NRATE term using separate runs resulting in the quasi-Type II hypotheses. Otherwise, the order is as listed. Note also that in this version of PROC MIXED, the denominator degrees of freedom for the spatial factors are not reduced by the between-subjects degrees of freedom.

The least-squares means were used to understand some of the effects highlighted in the previous section. We used adjusted means due to the unbalanced nature of yields over the fruiting locations on a plant. We did not consider absence of yield at a location on a plot to be zero yield. This component will be modelled in a separate analysis. Plots of the interaction least-squares means certainly indicate the presence of interactions but the general main effect trends seem to hold (Figures 2 and 3). It is easy to see why there is a significant interaction lack-of-fit from fitting linear and quadratic polynomials as the NRATE effect is not well modelled by these terms (Figures 2 and 4). The problems appear particularly important at the 5th, 6th, and 7th NRATE levels corresponding with 112, 140, and 168 kg ha^{-1}. These highest levels of nitrogen may be changing the way the fruiting on the plant responds with increasing nitrogen. See Boquet et al. (1993) for a discussion of these effects. The main effect of NODE appears suitable for modelling with simple polynomials and supports the fixed effects tests reported earlier (Figure 5).

5. Discussion

The most immediate result that we get from the model selection analysis is that any model of independence is not good. These models have zeros in off-diagonal positions in Σ while conditional independence models have zeros in off-diagonal positions in Σ^{-1}. The best fitting model using the AIC criterion was the variable-order ante-dependence model. Note that AIC, however, does not impose as strong of a penalty as BIC. When BIC is used the best fitting models appear to be the linear with nugget and Gaussian with nugget spatial covariance structures. However, using BIC, all spatial structures with a nugget fit about equally well and the variable-order ante-dependence model also fit well. Using AIC all ante-dependence models with order larger than zero fit reasonably well. Also note that the autoregressive model does not fit especially well and may suggest that a heterogeneous structure might be more appropriate. A test comparing two models using AIC is possible (see Linhart 1988), but requires knowledge of the asymptotic covariance matrix of parameter estimates. We used the AIC and BIC criteria as measures of fit as opposed to tests, per se. We believe that additional criteria should also be used. The variogram structures and conditional independence structures each imply different aspects of the biology of the fruiting of cotton and so selection of a model should also consider the biological implications of the model. These considerations will be explored further in later work.

Some interesting information derived from the variable-order ante-dependence model is the result that low order nodes (low on the plant) have very local dependencies, while nodes located higher up on the plant have dependencies all the way to the bottom or near the bottom of the plant. We will investigate this result further in terms of the physiology of the cotton plant. The dependencies may be related to the way in which the cotton plant adds bolls as it grows. Note that the spatial models, although they appear to fit well, make an assumption of a covariance process that does not depend upon the node but depends only upon the distance between nodes (stationarity). Thus, if the dependencies change with the location, the ante-dependence models might be more appropriate and would certainly be much more flexible. More work is needed here to determine whether or not the

variable-order ante-dependence models are fitting signal or noise. For the spatial models, we have assumed that the distance between nodes is given by the difference between the node numbers. Distances might better be determined in terms of the plant's physiology. An advantage of the ante-dependence model is that it uses only the orderings of the nodes and does not require that the distance between nodes be explicitly determined. Lastly, we have not included the effects of the years of sampling since only 1 year was analyzed for this paper. In addition, the REP effect could be handled as a random component in a mixed models approach. The more interesting addition, however, will be adapting these methods to handle 2 repeated measures factors, the NODE and the fruiting sites within nodes.

Our goal in using spatial models for the cotton data are several. First, depending upon how far the traditional split-plot models deviate from the compound-symmetry structure, spatial models may be much more efficient and provide more powerful tests of hypotheses of interest. Second, the spatial models may provide insight into the biology or physiology of cotton fruiting by describing the relationships among fruiting sites. Lastly, spatial models can be used to make predictions of yields at fruiting locations, where the predictions exploit knowledge of yields at lower order nodes. This could be very useful in growth and yield models for cotton plants during the growing season.

6. Summary

Geostatistical spatial models were fit to repeated measures data of within-plant boll average boll weights, where the spatial locations were taken as the sympodial branch NODE along the stem. Generally these models fit well, particularly when compared with standard approaches that might use an independence or split-plot model. The Gaussian spatial structure including a nugget fit well according to the AIC and BIC information criteria. Ante-dependence models were also fit to the same data. These models exploit the ordering of the nodes along the plant but make no requirement about their spatial coordinates. Fixed-order ante-dependence models did not fit as well as the spatial models, due to the larger number of parameters in the models. However, the variable order ante-dependence model had the smallest AIC value of all models and also had a competitive BIC value. This model indicated that low-order nodes were dependent at most upon previous nodes, while nodes much higher on the stem (Node 14 and above) were dependent upon most of the lower-order nodes. Thus, the variable order ante-dependence model accounts much better for the growth and physiology of the plant than the isotropic spatial models.

7. References

Boquet, D.J., E.B. Moser, and G.A. Breitenbeck. 1993. Nitrogen effects on boll production of field-grown cotton. *Agronomy Journal* 85, 34-39.

Crowder, M.J. and D.J. Hand. 1990. *Analysis of Repeated Measures*. Chapman and Hall, London.

Lindsey, J.K. 1993. *Models for Repeated Measurements*. Clarendon Press, Oxford.

Linhart, H. 1988. A test whether two AIC's differ significantly. *South African Statist. J.* 22, 153-161.

Macchiavelli, R.E. 1992. *Likelihood-based Procedures and Order Selection in Higher Order ante-dependence Models.* Ph.D. Thesis, The Pennsylvania State University.

Macchiavelli, R.E., E.B. Moser, and S.R. Pezeshki. 1993. Use of antedependence models in forestry trials. Department of Experimental Statistics, LSU, *Research Report Series* No. RR-93-25.

Macchiavelli, R.E. and S.F. Arnold. 1994. Variable order ante-dependence models. *Communications in Statistics A* 23(9), 2683-2699.

Macchiavelli, R.E. and E.B. Moser. 1994. Analysis of repeated measurements with ante-dependence models. Department of Experimental Statistics, LSU, *Research Report Series* No. RR-94-31.

Moser, E.B., A.M. Saxton, and S.R. Pezeshki. 1990. Repeated measures analysis of variance: application to tree research. *Can. J. For. Res.* 20, 524-535.

Munroe, J.M. and H.G. Farbrother. 1969. Composite plant diagrams in cotton. *Cotton Grw. Rev.* 46, 261-282.

SAS Institute. 1989. *SAS/IML Software: Usage and Reference, Version 6, First Edition.* SAS Institute, Inc., Cary, NC.

SAS Institute. 1992. The Mixed procedure. Pages 287-366 in *SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07.* SAS Institute, Inc., Cary, NC.

TABLE 1. Correlations (upper triangle) and partial correlations given intermediate nodes (lower triangle) of residuals computed from the multivariate model imposing no covariance restrictions.

Node	8	9	10	11	12	13	14	15	16	17	18
8	1.00	0.74	0.57	0.34	0.33	0.20	0.55	0.08	0.06	0.13	0.11
9	0.74	1.00	0.67	0.49	0.21	0.31	0.21	-0.19	0.10	-0.11	-0.02
10	0.15	0.67	1.00	0.52	0.11	0.15	0.24	-0.09	0.02	0.15	0.15
11	-0.07	0.22	0.52	1.00	0.32	0.45	0.16	0.12	0.39	0.28	0.20
12	0.31	0.12	-0.07	0.32	1.00	0.32	0.48	0.44	0.47	0.55	0.31
13	-0.06	0.20	-0.11	0.39	0.32	1.00	0.40	0.60	0.61	0.32	0.50
14	0.61	-0.09	0.31	-0.13	0.41	0.40	1.00	0.51	0.43	0.61	0.52
15	0.34	-0.53	-0.23	-0.28	0.20	0.50	0.51	1.00	0.46	0.68	0.72
16	-0.28	-0.16	-0.25	0.11	0.29	0.45	0.26	0.46	1.00	0.52	0.31
17	-0.23	-0.31	0.15	0.36	0.21	-0.41	0.37	0.59	0.52	1.00	0.67
18	-0.27	0.02	0.10	0.00	-0.17	0.33	0.16	0.50	-0.07	0.68	1.00

TABLE 2. The Cholesky decomposition A of the rescaled concentration matrix of residuals computed from the multivariate model imposing no covariance restrictions. Multiply the matrix $A'A$ by 1.73 to get Σ^{-1}.

| Node	8	9	10	11	12	13	14	15	16	17	18
8	1.00										
9	-0.44	0.44									
10	-0.01	-0.31	0.46								
11	-0.09	-0.05	-0.23	0.41							
12	0.00	-0.15	0.02	-0.08	0.39						
13	0.20	-0.17	-0.04	-0.25	-0.02	0.45					
14	-0.29	-0.08	-0.10	0.06	-0.06	-0.04	0.30				
15	-0.20	0.11	0.11	0.15	-0.03	-0.14	-0.01	0.31			
16	0.08	0.01	0.07	0.00	-0.03	-0.16	-0.04	-0.05	0.21		
17	0.06	0.14	-0.04	-0.15	-0.11	0.20	-0.11	-0.10	-0.15	0.30	
18	0.09	-0.00	-0.03	-0.00	0.05	-0.10	-0.05	-0.14	0.02	-0.17	0.19

TABLE 3. Model selection results for the various covariance structures using the Akaike's AIC and Schwarz's BIC information criteria. NC indicates that PROC MIXED did not converge and *NA* in the TYPE column indicates that the covariance structure is not available in PROC MIXED.

| Covariance Structure	TYPE	d_k	AIC	BIC
Unstructured	UN	66	-810.0	-858.4
Simple (Independence)	SIM	1	-876.6	-877.3
Compound Symmetry	CS	2	-845.9	-847.4
Autoregressive Order 1	AR(1)	2	-836.8	-838.2
Ante-dependence Order 0	UN(1)	11	-868.1	-876.2
Ante-dependence Order 1	NA	21	-826.4	-841.8
Ante-dependence Order 2	NA	30	-818.5	-840.5
Ante-dependence Order 6	NA	56	-807.6	-848.6
Ante-dependence Variable Order Orders={0 1 1 1 1 2 6 7 4 6 3}	NA	43	-799.9	-831.4
Spatial Linear	SP(LIN)	2	-871.7	-873.2
Spatial Linear with Nugget	SP(LIN)	3	-823.6	-825.8
Spatial Linear log	SP(LINL)	2	NC	NC
Spatial Linear log with Nugget	SP(LINL)	3	-827.1	-829.3
Spatial Exponential	SP(EXP)	2	-836.8	-838.3
Spatial Exponential with Nugget	SP(EXP)	3	-825.3	-827.5
Spatial Gaussian	SP(GAU)	2	-850.7	-852.2
Spatial Gaussian with Nugget	SP(GAU)	3	-823.5	-825.7
Spatial Spherical	SP(SPH)	2	-877.6	-879.0
Spatial Spherical with Nugget	SP(SPH)	3	-824.2	-826.4

TABLE 4. Fixed-effects tests of hypotheses. NDF and DDF are numerator and denominator degrees of freedom, respectively. The sidedress NRATE level has been removed from the NRATE linear, quadratic, and lack of fit tests.

| Source	NDF	DDF	Type I F	Type III F	Pr > F
REP	3	21		20.79	0.0000
NRATE	7	21		3.46	0.0347
Linear	1	21	0.11		0.7412
Quadratic	1	21	9.81		0.0019
Lack of fit	4	21	3.30		0.0117
NODE	10	255		59.78	0.0000
Linear	1	255	521.04		0.0000
Quadratic	1	255	62.89		0.0000
Lack of fit	8	255	1.06		0.3946
NRATE*NODE	69	255		1.93	0.0001
Linear by Linear	1	255	21.69		0.0000
Linear by Quadratic	1	255	0.00		0.9814
Quadratic by Linear	1	255	0.03		0.8729
Quadratic by Quadratic	1	255	0.00		0.9711
Lack of fit	63	255	1.68		0.0027

FIGURE 1. Empirical semi-variogram constructed from the residuals from the completely unstructured covariance model. Note that the fixed effects model components have been removed (de-trended).
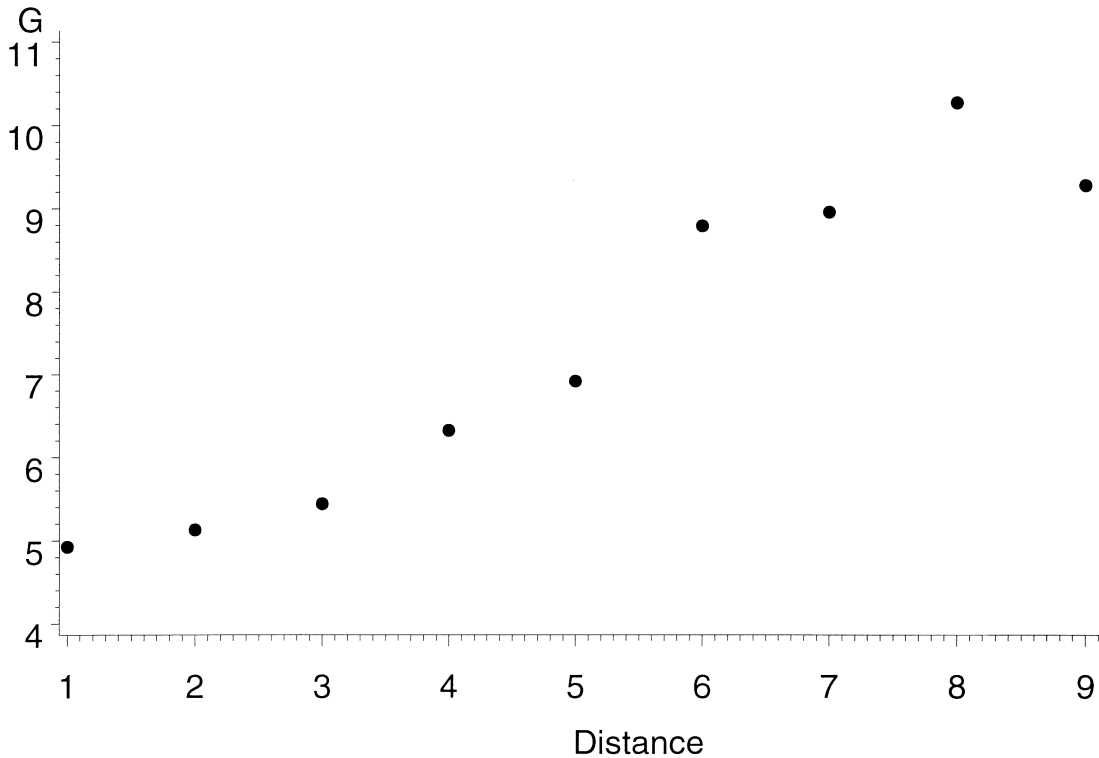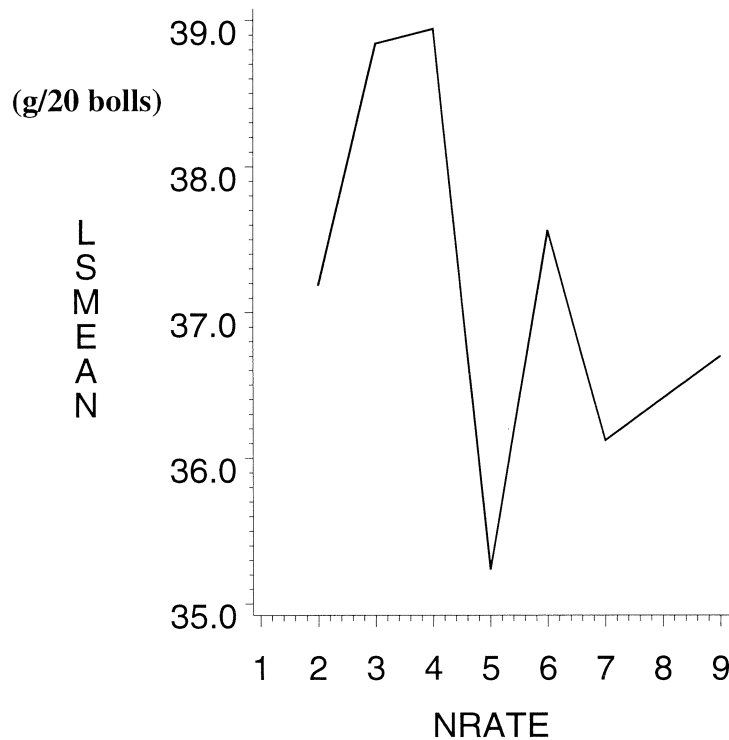

FIGURE 2. Interaction of NRATE with NODE as plotted using least-squares boll weight means. See the text for NRATE levels.
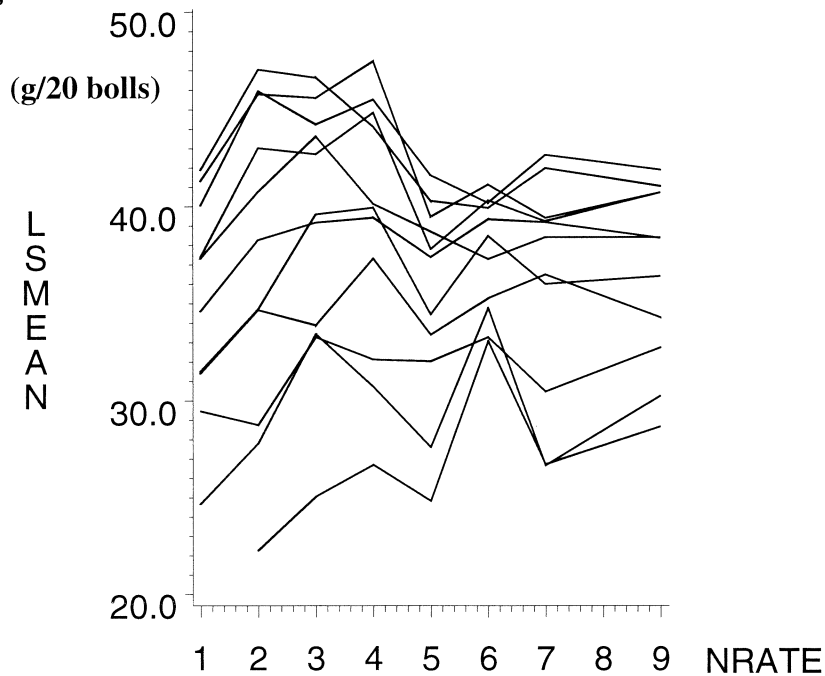
FIGURE 3. Interaction of NRATE with NODE as plotted using least-squares boll weight means. NRATE levels are the different lines.
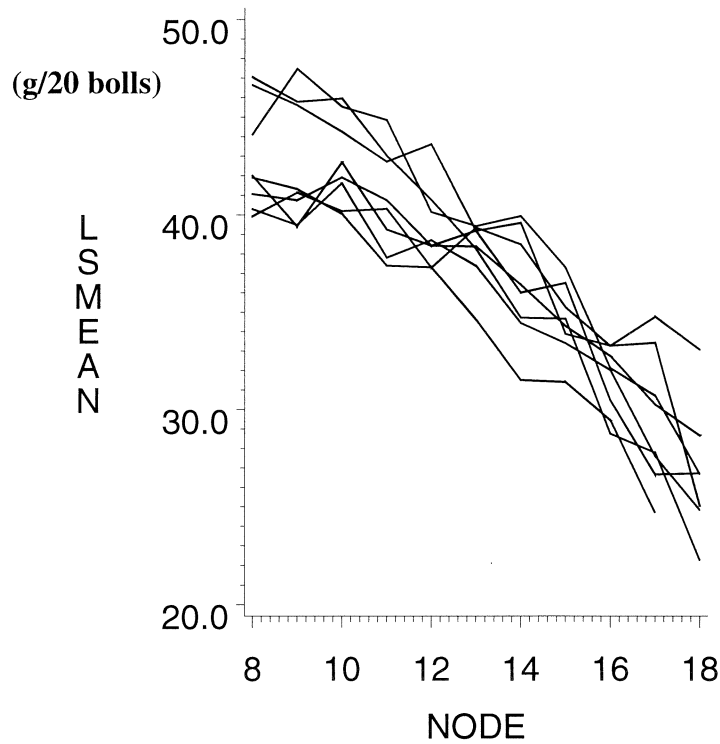
FIGURE 4. Main effect of NRATE plotted as least-squares boll weight means. See the text for NRATE levels.

FIGURE 5. Main effect of NODE plotted using the least-squares boll weight means.