Evidence for conformal invariance of crop yields

BY PETER MCCULLAGH* AND DAVID CLIFFORD

Department of Statistics, University of Chicago, 5734 South University Avenue, Chicago, IL 60637, USA

The aim of this paper is to study the nature of spatial correlation of yields of agricultural crops. The focus is primarily on natural or non-anthropogenic spatial variation, patterns that cannot be explained by topography, by variety or treatment effects, or by agricultural practices. Conformal invariance implies stationarity and isotropy, and also determines the rate of decay of spatial correlations. The resulting Gaussian model is studied empirically to see whether it describes satisfactorily the pattern of spatial correlations observed in field trials of various crops. By embedding the law in a larger statistical model, a convolution of white noise and the Matérn class having a range parameter λ^{-1} and a smoothness parameter ν, and by gathering data of sufficient range and quantity, the model predictions were tested. Twenty-five examples of crop yields are studied, including cereals, root crops and other vegetables, nut, citrus and alfalfa yields. At the scale of typical field trials, we find that non-anthropogenic variation is reasonably close to isotropic. Furthermore, we find consistent evidence that the range parameter tends to be large and the smoothness parameter small. The large value of the range parameter confirms Fairfield Smith (Fairfield Smith 1938 *J. Agric. Sci.* **28**, 1–23), who found that spatial correlation in agricultural processes decreases with distance, but at a slower rate than exponential. The small value of the smoothness parameter means that, by Matérn standards, agricultural processes are rough. For each of the examples studied, the limiting model fits the data just as well as the full model, in reasonable agreement with the hypothesis of the conformal model that $(\lambda, \nu) = (0, 0)$ for all crops in all seasons.

Keywords: agricultural field trial; conformal transformation; contrast measure; Fairfield Smith's law; Gaussian random field; isotropy

1. Introduction

Spatial correlation in agricultural field trials has long been recognized. Mercer & Hall (1911) studied the variability of plot yields as a function of plot size, seeking to determine effective and economical plot sizes for field experiments. Fairfield Smith (1938) developed an empirical power law for this relationship, making it clear that spatial correlations do not decay fast with distance. The implications of spatial correlation for parameter estimation were apparent to Fisher, whose solution was based on blocking, randomization and an associated decomposition of the total sum of squares. Papadakis advocated 'local control' in the form of an

* Author for correspondence (pmcc@galton.uchicago.edu).

iterative algorithm (Papadakis 1937), in which each plot is judged according to the average of its neighbours. The potential for using explicit nearest neighbour models for the analysis of field experiments became apparent following Besag (1974). Bartlett (1978) showed that the Papadakis algorithm could be interpreted in terms of an explicit lattice Markov model on the plots, and this approach has been subsequently used and extended by Wilkinson *et al.* (1983), Besag & Kempton (1986) and Besag & Higdon (1999).

An alternative procedure is to construct planar processes as direct products of simpler one-dimensional processes (Cullis & Gleeson 1991; Gilmour *et al.* 1997). Such constructions are open to criticism on the grounds of their dependence on the spatial coordinate axes and the behaviour of linear predictors (Stein 1999, §2.11). As a practical matter, however, they are quite effective for rectangular field layouts.

In the geostatistical literature, the approach to spatial modelling has been dominated by the use of random fields, almost invariably Gaussian random fields (Journel & Huijbregts 1978; Cressie 1993). These are defined continuously on the plane, and are not tied to the usual rectangular lattice pattern of plots in field trials. For any particular arrangement of plots in the field, the joint distribution is obtained by integrating the covariance function over plots (Zimmermann & Harville 1991), so the parameters have an interpretation in terms of distance and area rather than neighbour coefficients. This paper follows the geostatistical route, beginning with a particular family of Gaussian random fields, a convolution of white noise and the Matérn class (Matérn 1986). The key process in this class, the de Wijs process, is also an infinitesimal isotropic limit of the lattice Markov model used by Besag & Higdon (1999) (see Besag & Mondal 2005).

This paper follows on from Fairfield Smith (1938) in the sense that emphasis is placed on the nature of spatial correlation of crop yields rather than on strategies for the estimation of treatment or variety effects.

2. Spatial models

(*a*) *Covariance function*

Spatial dependence is modelled as a planar Gaussian process, in which the covariance function or generalized covariance function at x, x' in the plane has the form

$$\sigma_0^2 \delta_{x-x'} + \sigma_1^2 K(|x-x'|), \tag{2.1}$$

with non-negative coefficients σ_0^2, σ_1^2. The Dirac delta function $\delta_{x-x'}$ is associated with stationary white noise, independent on non-overlapping sets. The second term is a stationary isotropic covariance function of the Matérn class,

$$K(t) = (\lambda t)^{\nu} K_{\nu}(\lambda t),$$

where K_{ν} is the Bessel function of order $\nu \geq 0$ (Matérn 1986; Stein 1999). In all cases, distance is measured in metres, so λ^{-1} is the range in metres. The family of covariance functions is thus indexed by four non-negative parameters, two variance components, a range parameter and a smoothness parameter.

The planar Matérn process with parameter (λ, ν) has a spectrum whose density at frequency ω is proportional to $(\omega^2 + \lambda^2)^{-\nu-1}$ (Stein 1999, p. 49). If the

range is finite, i.e. $\lambda > 0$, the spectrum is finite at the origin. Otherwise, for $\lambda = 0$, the spectrum is singular at the origin, and these limiting processes are said to have infinite range.

The exponential model with covariance function $\exp(-\lambda|x-x'|)$ corresponds to $\nu = 0.5$ in the Matérn class. The model with $\nu = 1$ was derived by Whittle (1954) as the solution of a stochastic partial differential equation. The power covariance function $-|x-x'|^{2\nu}$ is obtained in the limit as $\lambda \to 0$ for non-integer $\nu > 0$. The limiting process as $\nu \to 0$ and $\lambda \to 0$, with generalized covariance function $-\log|x-x'|$ on contrasts, is called the de Wijs process (de Wijs 1951, 1953), and the conformal model is the de Wijs process plus white noise in (2.1). The limiting process as $\nu \to 1$ and $\lambda \to 0$, with generalized covariance function $|x-x'|^2\log|x-x'|$, is called the thin-plate smoothing spline (Wahba 1990). Each power covariance function determines a generalized process defined on contrasts only.

One interpretation of (2.1) is that the white-noise term corresponds to measurement error and the second term corresponds to fertility effects in the soil (Besag & Kempton 1986; Besag & Higdon 1999). However natural this interpretation might appear, the model does not imply that the pattern of low- and high-yielding plots is persistent from one season to the next. On this point Mercer & Hall (1911, p. 112) remark that '(t)he variations … are determined by so many incalculable factors that they may fairly be described as casual, and a plot above average in one year is as likely as not to be below the average in the next year'. The same sentiment is echoed by Pearce (1978) in the discussion of Bartlett (1978). While persistent patterns may well be present, it is far from clear that these are the dominant source of plot-to-plot variation. The 1924 and 1925 alfalfa yields reported in §4(*f*) are positively correlated ($r = 0.65$), so the pattern of yields has a persistent component. By contrast, the correlations between the yields of straw and grain in 1925 and the yields of roots and tops in 1926 in Sawyer field are in the range $(-0.2, -0.3)$, a statistically significant negative correlation that might be explained by weather differences in the two years.

In the sense that it covers a wide range of processes, from smooth to rough and exhibiting long- and short-range spatial correlation, the spatial model (2.1) seems eminently suitable for applied work. However, it will be seen that agricultural processes are seldom smooth, so the model may be flexible in ways that are unhelpful. The model is also restrictive in a number of respects, the most obvious being that each of the processes is stationary and isotropic. In practice, there may be non-isotropic effects that have not been accommodated, so it is essential to have a routine test for isotropy. If anisotropy is found, it may be necessary to extend the model either by defining distance in an anisotropic way, or by the inclusion of anisotropic terms either in the mean or in the covariance function. An additional covariance term K' constant on rows may be helpful in circumstances where there is excess dependence within rows. In particular, additive independent row effects may be incorporated by defining K' to be the block factor or equivalence matrix for rows.

(*b*) *Covariance matrix*

Consider a design or field layout in which the yield is observed on n non-overlapping plots. Denote by $Y(A)$ the yield on plot A. According to the model

(2.1), the covariance of the yields on two plots is given by

$$\operatorname{cov}(Y(A), Y(A')) = \sigma_0^2 |A| \delta_{A,A'} + \sigma_1^2 \int_{A \times A'} K(|x - x'|) \mathrm{d}x \, \mathrm{d}x'$$

$$= \sigma_0^2 |A| \delta_{A,A'} + \sigma_1^2 (|A| \times |A'|) \operatorname{ave}_{A \times A'} K(|x - x'|),$$

where $|A|$ denotes plot area and $\mathrm{d}x$ is planar Lebesgue measure. Usually, all of the plots have the same area, in which case the covariance matrix for the vector Y is

$$\Sigma = \sigma_0^2 |A| I_n + \sigma_1^2 |A|^2 V, \tag{2.2}$$

where V_{ij} is the average of $K(|x - x'|)$, with x in plot i and x' in plot j. For the Matérn class, the elements in V depend on (λ, ν). Analytic expressions for the average are available in a few special cases, for example for rectangular plots and logarithmic or power covariance functions. Otherwise, the average is computed numerically for each parameter point by reducing the calculation to a one-dimensional integral (Clifford 2005).

The covariance matrix Σ in (2.2) depends on the physical dimensions of the plots and on the geometric arrangement of the plots in the field. The model makes no assumptions about the geometry, but our software is designed for congruent rectangular plots arranged in a lattice with perpendicular rows and columns. Interspersed fallow plots or missing values can be accommodated. To determine Σ, it is sufficient to give the plot dimensions plus the vectors of row and column spacings. In principle, this information should be readily available, but unfortunately it is not always recorded for historical data.

An intrinsic or generalized process has a kernel $\mathcal{K}$, a vector subspace consisting of certain functions on the plane, usually a low-dimensional subspace consisting of the constant functions plus low degree polynomials (Cressie 1993; Chilès & Delfiner 1999). A process with kernel $\mathcal{K}$ is defined on the domain $\mathcal{K}^0$, the space of linear functionals taking the value zero on the kernel. This domain is closed under smooth invertible planar transformation. When $\mathcal{K} = \mathbf{1}$ is the one-dimensional space of constant functions, the domain $\mathbf{1}^0$ is the set of planar contrasts. For a statistical model consisting of a set of processes, the domain of the model is the intersection of the domains of the individual processes. Since the limiting conformal model is defined on integrated contrasts, comparisons with the Matérn class are based on contrasts.

(c) Likelihood

The linear Gaussian model has a mean vector $\mu \in \mathcal{X} \subset \mathcal{R}^n$, a subspace generated by specific non-isotropic effects and covariance matrix Σ. The emphasis in the analyses that follow is on the parameters determining Σ, not the parameters determining μ. However, estimates of variety and other additive effects are obtained automatically as a by-product of these computations. If the model has a kernel $\mathcal{K}$, then necessarily $\mathcal{K} \subset \mathcal{X}$. Since row and column effects may be expected on account of mechanical tillage and harvesting, additive row and column effects may be included in $\mathcal{X}$ as a precautionary measure. An instance of such a pattern generated by tractor wheelings is described by Besag & Kempton (1986, p. 244). In many circumstances, it may be more economical to model additive row and

column effects by additive terms in the covariance function as described in §7. However, in the interests of uniformity for the primary analysis in §5, we prefer to err on the side of caution by eliminating additive effects by linear projection.

Let the columns of the $n \times p$ matrix X be a basis for $\mathcal{X}$, and let $Q = I - X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}$. For the model (2.1), the standardized residual vector $Q Y / \|QY\|$ has a distribution depending on $\sigma_1^2/\sigma_0^2, \lambda, \nu$ and independent of μ. The log-likelihood function based on this statistic is

$$\check{l}(\Sigma; y) = -\frac{n-p}{2}\log(y^{\mathrm{T}}\Sigma^{-1}Q_\Sigma y) - \tfrac{1}{2}\log|\Sigma| - \tfrac{1}{2}\log|X^{\mathrm{T}}\Sigma^{-1}X|,$$

where $Q_\Sigma = I - X(X^{\mathrm{T}}\Sigma^{-1}X)^{-1}X^{\mathrm{T}}\Sigma^{-1}$ (Tunnicliffe-Wilson 1989; Cruddas *et al.* 1989). Coincidentally, this is also the residual maximum likelihood (REML) profile log likelihood (Patterson & Thompson 1971; Harville 1977; Gleeson & Cullis 1987; Diggle 1988; Jiang 1996). Since the residual configuration statistic is dimensionless, the marginal likelihood function is unaffected by linear transformation of the response. It is thus irrelevant whether yield is measured in grams or pounds.

If the response is yield, σ_0^2 has the physical dimension $[\mathrm{M}^2][\mathrm{L}^{-2}]$ of squared mass per unit area, σ_1^2 has the physical dimension $[\mathrm{M}^2][\mathrm{L}^{-4}]$ and the variance ratio σ_1^2/σ_0^2 has physical dimension $[\mathrm{L}^{-2}]$. The inverse ratio is an area, and it is sometimes convenient in graphs to set the unit of area so that the variance ratio is approximately 1. Likewise, λ has the physical dimension $[\mathrm{L}^{-1}]$. Original units are used in the discussion of historical examples, but these are converted to metres in likelihood calculations.

(d) Test statistics

Let $Y \sim N(X\beta, \Sigma)$, and let $T = Y^{\mathrm{T}}AY$ be a quadratic form in the residuals, meaning that A is symmetric and $AX = 0$. To use T as a test statistic, we must have a reference distribution, exact or approximate. In the calculations that follow, we proceed as if Σ were known up to a multiplicative constant, so we require the distribution of T conditional on the weighted residual sum of squares.

Given that $Y^{\mathrm{T}}\Sigma^{-1}Q_\Sigma Y = n - p$, the distribution of $Q_\Sigma Y$ is uniform on the unit sphere in the relevant $(n-p)$-dimensional space. The conditional mean and variance of T are obtained in the form

$$E(T) = \mathrm{tr}(A\Sigma),$$

$$\mathrm{var}(T) = 2((n-p)\mathrm{tr}(A\Sigma)^2) - \mathrm{tr}^2(A\Sigma))/(n-p+2).$$

The conventional estimate of Σ is such that $Y^{\mathrm{T}}\hat{\Sigma}^{-1}Q\hat{\Sigma}Y = n - p$, so we use the formulae above with Σ replaced by $\hat{\Sigma}$ to obtain approximate moments for the test statistic.

If T is positive semi-definite, the standardized log ratio is $\log(T/E(T)) E(T)/\sqrt{\mathrm{var}(T)}$. Otherwise, the standardization is by subtraction of the mean and division by the standard deviation.

The suggested test for anisotropy is a score test, a pair of quadratic forms in the residuals, $Y^{\mathrm{T}}\Sigma^{-1}Q_\Sigma A\Sigma^{-1}Q_\Sigma Y$, one in which $A_{ij} = \cos(2\theta_{ij})$ and one in which $A_{ij} = \sin(2\theta_{ij})$. Here θ_{ij} is the direction of plot j from plot i, computed for convenience from the plot centres. A large absolute value of the first component

indicates anisotropy oriented along the rows and columns: the second component is most sensitive to anisotropy at $45°$ relative to the axes. These could be combined into a single statistic, but it is usually best not to do this because information about the direction of the anisotropy would thereby be lost.

3. Spherical and planar processes

Every planar process determines a process on the sphere S by stereographic projection through the north pole or antipodal point. If the planar process is defined on planar contrasts, the spherical process is defined on spherical contrasts. A planar contrast ρ is a measure or distribution, such that $\rho(\mathcal{R}^2) = 0$, and the value of Y at ρ is formally an integral against ρ. Typically, ρ has a density that is a positive constant on one or more plots and a negative constant on another set of plots, so $Y(\rho)$ is a difference of spatial averages. If $g : \mathcal{R}^2 \to S$ is invertible, the transformed distribution ρg^{-1} is a spherical contrast $(\rho g^{-1})(A) = \rho(g^{-1}A)$, for $A \subset S$ and $g^{-1}A \subset \mathcal{R}^2$. For a planar process Y defined on contrasts, the associated spherical process Y^g is $Y^g(\rho g^{-1}) = Y(\rho)$, defined on spherical contrasts.

A planar process that is stationary and isotropic does not determine a spherical process that is invariant under rigid spherical motions unless the planar process is also conformally invariant. The group of conformal planar transformations is generated by rigid planar motions plus rigid spherical motions. It consists of the planar transformations $x \mapsto g(x)$,

$$g(x) = \frac{ax + b}{cx + d}, \quad g^{-1}(x) = \frac{dx - b}{a - cx},$$

on the extended complex plane. Here, $x = x_1 + \mathrm{i}x_2$ is a complex number, and the coefficients a, b, c, d are complex numbers, such that $ad - bc \neq 0$.

A second, and quite different, argument for conformal invariance runs as follows. First, any transformation that preserves Euclidean properties of the space should have no effect on the process. Second, only local properties of Euclidean space matter. The first condition implies stationarity and isotropy, and also self-similarity if scalar dilations are included. Together, the two conditions imply conformal invariance.

The first argument for conformal invariance is the same as the argument for stationarity and isotropy, except that the argument is applied simultaneously to the plane and the sphere. On their own, arguments of this sort based on algebraic invariance are not compelling. Their relevance to agricultural processes can only be demonstrated by critical examination of data from suitable experiments.

A zero-mean Gaussian process defined pointwise is conformally invariant if $K(x, x') = K(gx, gx')$ for each transformation g in the group. For each $x \neq x'$, there exists a group element, such that $gx = 0$ and $gx' = 1$, so the process is conformally invariant only if $K(x, x') = \sigma^2$ for $x = x'$ and $\rho\sigma^2$ for $x \neq x'$. In essence, this means that the process is a random constant.

Suppose, however, that the process is defined on contrasts and K is defined pointwise in such a way that

$$\mathrm{cov}(Y(x_1) - Y(x_2), Y(x_3) - Y(x_4)) = K(x_1, x_3) + K(x_2, x_4) - K(x_1, x_4) - K(x_2, x_3).$$

The process is conformally invariant if for each group element g

$$\mathrm{cov}(Y(gx_1) - Y(gx_2),\, Y(gx_3) - Y(gx_4)) = \mathrm{cov}(Y(x_1) - Y(x_2),\, Y(x_3) - Y(x_4)),$$

which implies that

$$K(gx_1, gx_3) + K(gx_2, gx_4) - K(gx_1, gx_4) - K(gx_2, gx_3)$$

$$= K(x_1, x_3) + K(x_2, x_4) - K(x_1, x_4) - K(x_2, x_3).$$

Given that the cross-ratio,

$$\frac{(x_1 - x_3)(x_2 - x_4)}{(x_1 - x_4)(x_2 - x_3)},$$

is invariant under the group, we observe that the invariance condition is satisfied by the real and imaginary parts of $K(x, x') = \log(x - x')$. Since the cross-ratio is maximal invariant, these are the only possibilities, and since K is real symmetric the only possibility is $\log|x - x'|$, which is the de Wijs process. The only remaining problem, that the variance of an atomic contrast $Y(x_1) - Y(x_2)$ is infinite, is resolved by restricting attention to non-atomic contrasts, usually contrasts having a density.

For further details concerning fractional linear transformation and stereographic projection, the reader is referred to Hille (1965, §3.1), Carathéodory (1954, §§31–35) or Rudin (1987, p. 278).

4. The data

The arguments leading to the conformal model (McCullagh 2002, p. 1307) are specific to planar processes, considered also as processes on the sphere by stereographic projection, but are not otherwise crop specific. In our search for examples suitable for testing the model, we tried to obtain data from a wide range of crops each grown under reasonably uniform conditions with plots of known size and spatial arrangement. Modern precision agriculture data collected directly by the harvester, with spatial position determined by the GPS satellite system, might seem ideal for this purpose. However, it quickly becomes apparent that the degree of mixing within the harvester, and the time taken for the grain to pass through, are such that the yield recorded at position x is a mixture from plants spread over many metres upstream (Whelan & McBratney 2002; Tørgersen & Waagepetersen 2004). The spatial correlation of lupin and cotton yields from precision agriculture experiments was studied by Clifford *et al.* (in press), and found to be approximately isotropic and conformally invariant after taking account of the transport and mixing. Historical data are less prone to spatial shifting and mixing, so the spatial distribution is easier to study.

Twenty-five examples of agricultural crops are considered, ranging from alfalfa, cereals, potatoes and brussels sprouts to lemons, oranges and walnuts. The first 21 of these are yields taken either from uniformity trials or from variety trials with few varieties. The remaining four examples are discussed in §7. In five cases, the yield is bivariate with both straw and grain for cereals or roots and

tops for beets. Plot area for cereal crops is highly variable, ranging from 0.5 square feet for the Fairfield Smith wheat trial, 15 square feet for the Wiebe trial, 90 square feet for the Mercer–Hall trial, to 4350 square feet for the Sawyer field trials. The examples cover a range of climates and soil types, from Hertfordshire to Montreal, from Canberra to California. Graphical model checks are described, mostly in connection with the conformal sub-model and only for the first few examples. Detailed comparisons with other Matérn sub-models are deferred to §5.

Complete geometric information is available for 12 of the examples, numbers 2–8 and 13–17, and the information for Sawyer field, examples 18–21, seems also to be reliable. Partial information is available for the remaining eight trials, so the missing geometric information must be inferred from the historical record or documented agricultural practices. The width of discard strips is the most common piece of missing information. Generally speaking, discard strips are small relative to plot dimensions, and in these cases the effect on likelihood calculations is not great. However, very large discard strips do sometimes occur, for example in the winter wheat experiment, or uneven discard strips in the Mercer–Hall experiment, and these can have a more substantial effect on likelihood calculations for spatial models. Exclusion of these eight trials from the summary analysis, or minor modification of the inferred geometric information, has little effect on the summary conclusions in §5. We have decided to retain all examples because some of these trials exhibit effects such as anisotropy, not seen to the same degree in the other trials.

(a) Example 1: spring barley

The first example is a uniformity trial for spring barley at the Plant Breeding Institute in Cambridge (Kempton & Howes 1981). The arrangement of plots in the field is a regular grid of 28 rows and seven columns, each plot being $5' \times 14'$. It is possible that discard strips $3'$–$6'$ might have been present, but the analysis here proceeds as if the plots were contiguous. The effect of discard strips is discussed briefly in §6.

This example is used to illustrate some of the details of the analyses and model checks, focusing particularly on the conformal model with $\lambda = \nu = 0$, and the thin-plate smoothing-spline model with $\lambda = 0$, $\nu = 1$. We take X to be the model matrix spanning the subspace of additive row and column effects, so these are automatically eliminated from the analysis. The thin-plate smoothing-spline process has a three-dimensional kernel consisting of the constant and the two coordinate functions, so $\mathcal{K} \subset \mathcal{X}$ as required. The log likelihood for each sub-model with fixed (λ, ν) depends on a single parameter, the variance ratio σ_1^2 / σ_0^2, or equivalently $\gamma = \sigma_1^2 / (\sigma_1^2 + \sigma_0^2)$ in the range [0,1]. Note that $\gamma = 0$ is the standard model with independent components, so the likelihoods are equal at this point. The evidence from the marginal log likelihoods in figure 1 for spatial correlation is overwhelmingly strong, and the plot suggests at first glance that there is little to choose between the two models. However, the maximum for the conformal model is 84.46, versus 81.69 for the thin-plate spline model. Admittedly, these models are not nested, but the difference of 2.77 in log-likelihood units is not negligible. The combined model, a convex combination of both covariance functions plus white noise, gives a maximized log likelihood of 86.24, so the
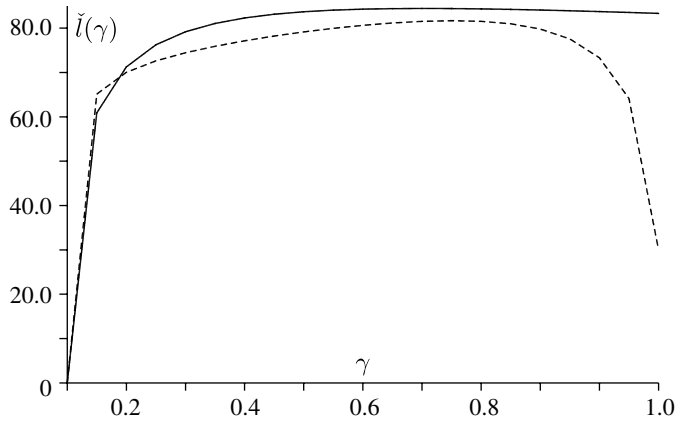
Figure 1. Marginal log likelihoods for two spatial covariance models: $-\log|x-x'|$ and $|x-x'|^2$ $\log|x-x'|/100$ (dashed line).

nominal 95% likelihood-based confidence set includes the conformal model and excludes the thin-plate spline. The overall maximum for the Matérn model (2.1) is 85.80 at $(\hat{\lambda}, \hat{\nu}) = (0.0, 0.32)$. The increase of 1.34 log-likelihood units on two degrees of freedom, corresponding to a nominal p-value of 26%, is well within the limits of sampling variation.

For each r, s, the contrast

$$D_{ij}^{rs} = Y_{ij} - Y_{i+r,j} - Y_{i,j+s} + Y_{i+r,j+s}$$

eliminates additive row and column effects and thus has zero mean. Its variance under the conformal model can be computed from the fitted variance matrix. The sum of squared contrasts,

$$T^{rs} = \sum_{ij} (D_{ij}^{rs})^2,$$

is a random quadratic form whose conditional mean and variance are given in §2d. The standardized log ratio is $\log(T^{rs}/\hat{T}^{rs})\hat{T}^{rs}/\hat{\sigma}$, where $\hat{T}^{rs}$ is the fitted mean and $\hat{\sigma}^2$ is the fitted variance.

For each r, s, the statistic T^{rs} was computed together with the mean and variance as determined by the fitted model. In figure 2a,c, the log ratio for the conformal model and its standardized version are plotted against the Manhattan metric $5r+14s$. Under the model, each standardized log ratio has mean zero and unit variance, but the individual points are not independent. Patterns are evident, some caused by dependence among the points, others by deviations from the model. Short- to moderate-range squared contrasts appear to have exactly the mean and variance that the model predicts. Long-range squared contrasts are, on the whole, a little larger than the model predicts. In fact, the deviations while possibly statistically significant are not at all large. To see what more typical plots looks like, the corresponding scatterplots are also displayed for the alternative model, in which $\log|x-x'|$ is replaced by $|x-x'|^2\log|x-x'|$. Five exceptionally large standardized ratios with values in excess of -4 have been suppressed in the second plot. This graphical test shows clearly that the second
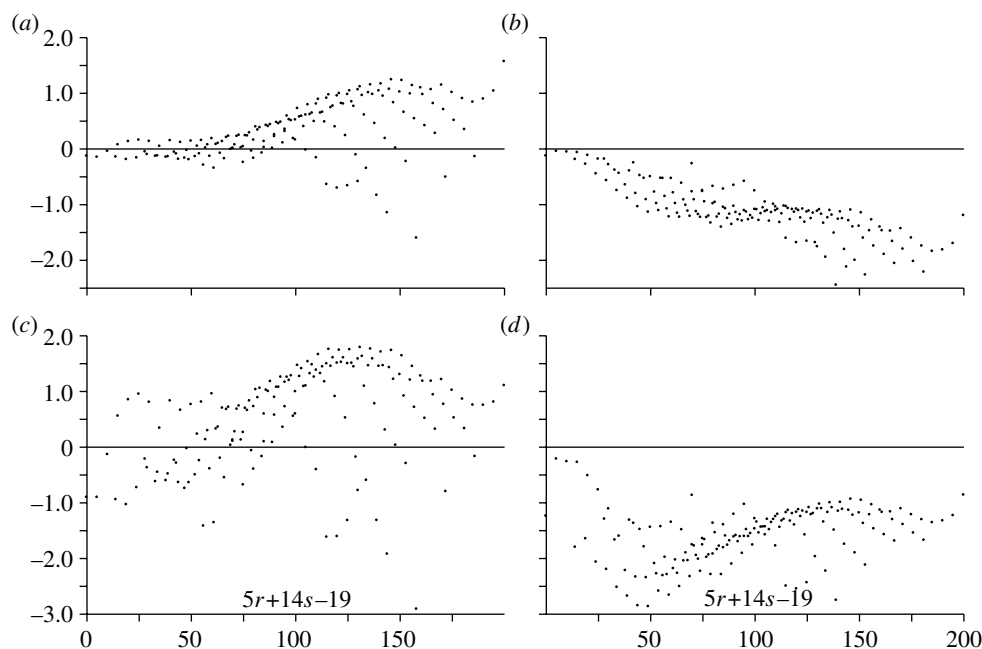
Figure 2. Log ratios $\log(T^{rs}/\hat{T}^{rs})$ of empirical variances to model-based estimates, unstandardized in the top row and standardized in the bottom row, plotted against the distance $5r+14s$. (a, c) The conformal model and (b, d) the alternative thin-plate spline model.

model overestimates the variances of most contrasts of the type considered here. To a much lesser extent, the conformal model errs in the other direction.

(b) Examples 2–8: fruit and nut plantations

Batchelor & Reed (1918) published extensive data on the yields of individual fruit trees taken from four citrus groves and one nut plantation in California, and one apple orchard in Utah. A brief description of all six datasets follows.

The first dataset consists of yields of individual 24-year-old trees from the navel-orange grove at Arlington Station in 1915–16. The grove consists of 20 rows from north to south, with 50 trees in a row, planted $22'$ by $22'$. In the analysis that follows, the grove was split into the eastern and western halves, consisting of 500 trees each in a grid 20×25. This procedure greatly reduces the computational effort, and also provides a check on the adequacy of the model.

The second dataset consists of yields from 495 ten-year-old navel orange trees in 1916 taken from a grove called Antelope Heights located at Naranjo. There are 15 rows of 33 trees, also planted $22' \times 22'$.

The third dataset consists of Valencia orange yields from 15-year-old trees at Villa Park in 1916. There are 12 rows of 20 trees, planted $21.5' \times 22.5'$. In the analysis reported here, the grid is taken as $22'$ square.

The fourth dataset consists of yields from 23-year-old Eureka lemon trees at Upland, California in 1916. There are 364 trees in 14 rows of 26 trees in each row, planted in a square grid of side $24'$.

The fifth dataset consists of yields from 280 walnut trees from a 24-year-old Santa Barbara softshell seedling grove. This planting was laid out 10 trees wide

and 28 trees long, on a square lattice of side $50'$. In the analysis reported here, each tree is assumed to occupy a square cell of side $25'$, so there is an additional $25'$ space between rows and columns.

The last dataset consists of yields of apples from a 10-year-old Jonathan apple orchard at Providence, Utah. There are eight rows of 28 trees, planted $16'$ apart, east and west, and $30'$ apart north and south. In the analysis reported here, each tree is assumed to occupy a cell of side $16'$, so there is a $14'$ vacant space in one direction.

In preparing the data for the computer, a few minor discrepancies were uncovered, where the values in the table did not tally with the published row and column totals. In some instances, these were corrected, though this could seldom be done with much certainty. In addition, Batchelor & Reed remark that the record for each tree known to be abnormal from disease was eliminated and replaced by the average of the eight surrounding trees. The text mentions 480 trees at Antelope heights and 240 trees in the Valencia grove, suggesting 15 imputed trees at Antelope heights and none for the Valencia grove. The imputed trees are not indicated, but a simple arithmetic check shows that the fraction must have been very small. Only two candidates exist in the Arlington grove, nine for Antelope heights, four for the Valencia grove, one each for Eureka and the walnut plantation, and none for the Utah apple orchard. However, further imputed trees may be concealed by rounding. Not all of these trees are necessarily imputed, so all trees are included in the analyses reported here.

In calculating and plotting the marginal log likelihood for γ in the conformal model, it was found to be convenient to take the unit of area to be 100 m^2. This choice has no substantive effect on the conclusions, but it does avoid extremely small and extremely large values of γ. For plots of this area, the two variance components are roughly equal, so γ is approximately one-half. In all cases, additive row and column effects have been eliminated from the analysis whether there are significant variations or not. The marginal log likelihoods for γ in the conformal model are displayed in figure 3 for each of the six plantations, with two parallel plots for the two halves of the Arlington grove. In the summary statistics listed in table 1, the log likelihoods measured relative to $\gamma = 0$ are given for the conformal model and the full model (2.1).

With the sole exception of the apple orchard at Providence, Utah, the evidence for spatial correlation is overwhelming. For example, the log-likelihood ratio statistic for the null hypothesis of no spatial correlation among the walnut trees is 2×5.17 on one degree of freedom for a p-value of 0.13%. For the Providence data, the log likelihood is fairly flat for small γ, so all that can be said is that the data are not consistent with values of γ in excess of about 0.4. None of this is surprising. Any reasonable spatial covariance model is sufficient to detect spatial correlation in these data. The main conclusion to be drawn is that the increase in log likelihood for the full Matérn model over the conformal sub-model is negligible in all cases. Standard asymptotic distribution theory suggests that $\Delta \tilde{l} \sim (1/2)\chi_2^2$, but this calculation ignores the fact that the conformal model is a boundary point in the parameter space. Regardless of regularity conditions, the combined increase of 4.73 log-likelihood units provides no suggestion of a departure from the conformal model.

It is quite remarkable that these six plantations, three of oranges, one of lemons, one of walnuts and one of apples, are fairly consistent with each other, so
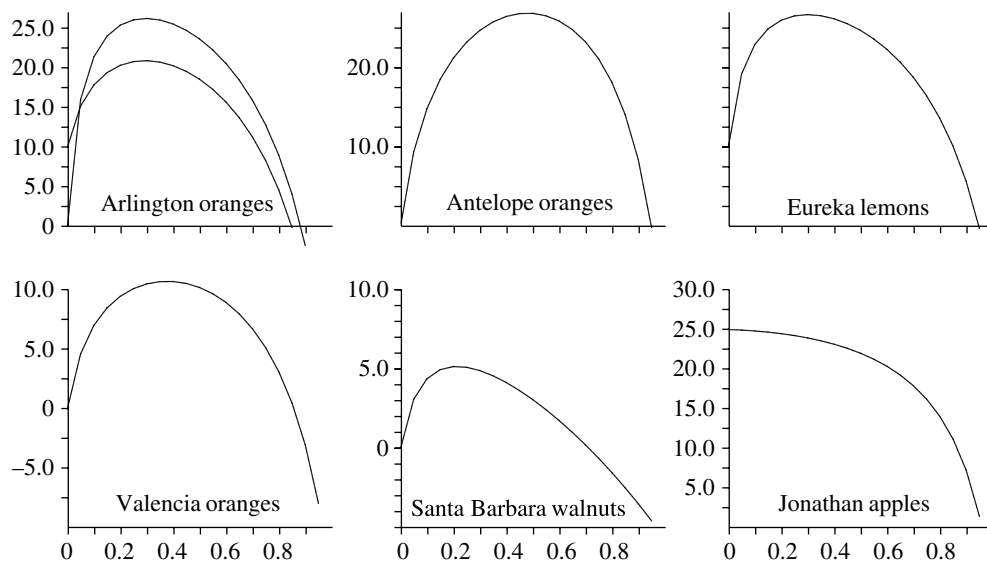
Figure 3. Marginal log-likelihood plots for γ in the conformal model applied to yields from six fruit and nut plantations.

Table 1. Summary fruit and nut statistics.

plantation	$\hat{\gamma}$	se($\hat{\gamma}$)	$\breve{l}(\hat{\gamma}, 0, 0)$	$\breve{l}(\hat{\gamma}, \hat{\lambda}, \hat{\nu})$	$\Delta \breve{l}$
Arlington I	0.297	0.082	26.19	26.21	0.02
Arlington II	0.294	0.091	10.89	12.44	1.55
Arlington I and II	0.341	0.062	74.85	74.85	0.00
Antelope	0.483	0.086	26.91	27.57	0.66
Eureka	0.302	0.091	16.71	17.58	0.87
Valencia	0.382	0.120	10.75	11.06	0.31
walnuts	0.213	0.103	5.17	6.15	0.98
Jonathan	0.000	—	0.00	0.34	0.34

far as the conformal coefficient γ is concerned. The estimated common value, obtained by maximizing the sum of the seven marginal log-likelihood functions, is $\hat{\gamma} = 0.327$, or $\hat{\gamma}/(1 - \hat{\gamma}) = 0.485$ per unit area of $100\,\mathrm{m}^2$, and the log likelihood at that point is 93.11. The likelihood ratio test statistic for the hypothesis of a common value is $2(96.61 - 93.11) = 6.97$ on six degrees of freedom. This rough and ready analysis treats the two halves of Arlington as independent which they are not. Even if the degrees of freedom are reduced to five to compensate, the evidence against a common value is not strong.

(c) Examples 9–10: Great Knott wheat

The Mercer & Hall data on wheat yields on Great Knott field at Rothamsted in the summer of 1910 have been used by later authors (Fairfield Smith 1938; Whittle 1956, 1962; Besag 1974; McBratney & Webster 1981; Wilkinson *et al.*

1983) for a wide range of purposes. Although most analyses focus on the grain yields, the data consist of both grain and straw yields on a grid of 20 rows, each 10.82′ wide, and 25 columns each 11 drills wide. Because of variations in drill width, it is hard to put a precise number on the column width, but it appears to be a little over 8′. Details of the experiment taken from the Mercer–Hall paper are reproduced in Andrews & Herzberg (1985) together with the data. One minor discrepancy in row 3, column 21 was noted and corrected.

To the best of our knowledge, all previous spatial analyses have proceeded as if the grid of plots were spatially regular. Certainly, it would be convenient if this were so, but Mercer & Hall's statement that …*a strip of varying breadth containing the furrow [was] left uncut between each row of plots* leaves little room for ambiguity. Discard strips were present in one direction, between the 25 rows in the field or columns on the page. Reading further, we find that the total width of discard strips between columns was around 160′. Even if the full implications of isotropy and the conformal model are accepted, the internal information available from the residuals about the spatial configuration of plots is not great. Nevertheless, we maximize the sum of the log likelihoods for grain and straw over the possible column configurations, subject to non-negative discards adding to 160′. Despite the paucity of information, the estimated discard widths in feet rounded to the nearest column width are rather striking.

$$16, 0, 0, 16^*, 0, 8, 16^*, 8, 0, 16^*, 0, 8, 8, 16^*, 16, 0, 0, 8^*, 8, 0, 0, 0^*, 16, 0.$$

Nearly half the estimated discards are zero or effectively zero, and on the whole, the larger estimated discards tend to coincide with the furrows marked by an asterisk. The discrepancy at the last furrow may be attributed to severe local anisotropy of the process, i.e. failure of the model, or to a misalignment in the recording of the furrow position. A very rough impression of the precision of the estimated discards may be inferred by comparing the configuration in which the 16′ discard at columns 23/24 is moved to the presumed furrow at columns 22/23. The decrease in log likelihood is 5.5 units for grain and 2.5 for straw. A uniform discard width of 6.67′ gives a decrease of 9.4 units for grain and 10.1 for straw. The information from grain and straw is fairly consistent regarding the inferred discards, and the evidence against uniformity of discards is fairly strong.

With freshly collected data, the breadth of all discards would be known, so we proceed as if the known values are those listed above. This is certainly not ideal, but on balance it seems preferable to proceeding as if the discards were of equal width. The conformal model achieves a marginal log likelihood of 52.22 for grain and 39.07 for straw, these being measured relative to the conventional white-noise model. The corresponding values for the full Matérn model (2.1) are 53.07 for grain and 41.04 for straw.

(*d*) *Examples 11–12: mangolds*

The data on mangold yields, also published by Mercer & Hall (1911), are taken from a uniformity trial at Rothamsted. The plants were grown on drills 2.4′ wide running west to east, each plot consisting of three drills 30.25′ long. Roots and leaves were weighed separately. The harvested area, one acre in area, contained 20 rows of three drills each, and 10 columns. Discard strips are not mentioned, but given the practice at the time we assume a discard of three drills.

Table 2. Log likelihoods for two anisotropic models.

	conformal		Matérn	
	isotropic	anisotropic	isotropic	anisotropic
roots	8.86	15.84	9.55	17.16
tops	5.61	7.10	7.46	8.66

Root crops such as beets, mangolds, turnips and swedes (rutabaga) are grown atop raised drills, so that the roots in different drills are physically more separated than they would be if the drills were not raised. The tops are also separated, but to a lesser extent depending on the height and width of the plant. Similar remarks apply to carrots, leeks, parsnips and potatoes, which are grown in the drills. For this reason, a degree of geometric anisotropy associated with the drills should come as no surprise. With additive row and column effects eliminated, the isotropic conformal model achieves a log likelihood of 8.86 for roots and 5.61 for leaves. The test for isotropy is $(-2.54, -0.45)$ for roots and $(-1.68, -0.46)$ for leaves, strongly indicating anisotropy in the direction of the drills. By introducing a geometric anisotropy factor, such that distance perpendicular to the drills is inflated by a factor of τ, the log likelihoods are increased to 15.84 at $\tau=5.5$ for roots and 7.10 at $\tau=2.3$ for leaves. The values for the full Matérn model are shown in table 2.

Given the assumed discard widths, the evidence for anisotropy is particularly strong for the roots, but the inflation factor is not well determined, and the data are consistent with a common value of 2 or above. The standard form of bivariate Gaussian model, in which the variance components in (2.2) are replaced by 2×2 matrices, requires a common value for the anisotropy parameter. Nevertheless, for the purposes of this project, roots and tops are analysed as separate processes with unconstrained anisotropy parameters. Although Mercer & Hall do not mention discard strips in connection with the mangolds, it is worth bearing in mind that the apparent anisotropy could equally well be explained by larger or irregular discard strips.

(e) Examples 13–15: three wheat trials

The winter wheat data are taken from Besag & Kempton (1986). The 364 plots are arranged in a regular grid of 52 rows and seven columns, each plot being 1.5 m × 4.5 m. Two varieties are grown in alternating plots, so X is the model matrix for additive row, column and variety effects. No information is given on discards, but this is not a uniformity trial, so discards must have been present. The analysis shown in table 3 assumes that the discard is 1 m wide on all sides. One peculiarity in these data is that the recorded plots are in fact check plots in a larger variety trial. Between each pair of columns there are five additional plots, so the total space between columns is $5 \times 4.5 + 1 = 23.5$ m.

If the unequal spacing of rows and columns is ignored, the maximized log likelihood for the conformal model is 31.88 compared with 36.47 for the actual configuration. The likelihood function provides clear evidence of this geometric

Table 3. Maximized marginal log likelihoods for six isotropic spatial models.

data	spatial covariance model					
	conformal	Matérn[a]	Bessel$_0$[a]	exp'l[a]	Whittle[a]	power[a]
spring barley	84.46	1.34	0.00	1.00	−0.14	1.34
Arlington I	26.19	0.02	0.00	−0.43	−1.10	0.02
Arlington II	10.89	1.55	1.55	1.50	1.46	1.43
Antelope	26.91	0.66	0.60	0.66	0.62	0.29
Eureka	16.71	0.87	0.00	0.84	0.70	0.87
Valencia	10.75	0.31	0.00	0.24	−0.12	0.31
walnuts	5.17	0.98	0.00	0.72	0.96	0.92
Jonathan	0.00	0.36	0.12	0.17	0.20	0.04
M-H grain	52.22	0.85	0.85	0.15	−0.32	1.21
M-H straw	39.07	1.97	1.94	1.95	1.88	1.46
mangold roots[b]	15.84	1.45	0.03	0.45	0.67	0.05
mangold tops[b]	7.10	1.64	0.73	1.00	1.16	0.19
winter wheat	36.47	0.15	0.15	−0.67	−1.40	0.06
Wiebe wheat	589.90	2.84	0.00	−1.08	−8.36	2.84
F-S wheat	12.02	0.69	0.68	0.69	0.67	0.64
alfalfa 1924	43.56	3.75	0.00	2.44	3.02	2.07
alfalfa 1925	22.34	3.75	0.00	1.92	2.77	2.09
Sawyer grain	11.41	1.47	0.00	0.91	0.70	0.97
Sawyer straw	10.50	1.20	0.00	0.66	0.79	0.67
Sawyer tops	5.26	1.14	0.00	0.64	0.99	1.08
Sawyer roots	8.24	0.26	0.00	0.11	−0.16	0.26
brussels sp	1.92	0.40	0.00	0.10	0.16	0.06
oats grain (§7*b*)	3.35	0.80	0.00	0.31	0.47	0.30
oats straw	3.16	0.72	0.00	0.28	0.41	0.17
potatoes (§7*c*)	7.59	0.39	0.19	0.38	0.39	0.04
total	1051.03	29.56	6.84	14.94	6.42	19.38

[a]Increment over the conformal model.
[b]With anisotropy parameter.

irregularity, and the anisotropy test is $(2.28, -0.53)$, correctly indicating anisotropy with row and column orientation.

Example 14 published by Wiebe (1935) is a uniformity trial of Federation wheat grown at Aberdeen Station, Idaho in 1927. For an alternative spatial analysis, see Zimmerman & Harville (1991). Each plot consists of a single row $15'$ long; each column consists of 125 plots $12''$ wide with no separation. Irrigation control levees separated most of the 12 columns, these being roughly $3'$ wide where present (Andrews & Herzberg 1985, p. 43). There are no discard strips.

Example 15 is a uniformity trial of Waratah wheat harvested at Canberra in December 1934, and included as the lead example by Fairfield Smith (1938) in his landmark empirical study of spatial variation. The plots in this experiment are very small, $6'' \times 12''$, with 1080 plots in a 30×36 grid with no discard strips. These data have been resurrected and are available from the Rothamsted library.

For the various cereal trials, the values of the variance ratio $\hat{\sigma}_0^2/\hat{\sigma}_1^2$ in square metres for the conformal model are as follows:

MH-g	WW	Wb	F-S	Saw-g	Sp-b	Oat-g	MH-s	Oat-s	Saw-s
10.33	12.92	1.36	0.65	5.18	0.73	250.86	11.55	256.22	15.25

The first seven of these are for grain, the last three for straw. The three straw–grain pairs have similar ratios, but other than that there appears to be no simple pattern.

(f) Examples 16–21

Summerby (1936) published the yields from a number of uniformity trials on various crops at MacDonald College, Montreal. Examples 16 and 17 are alfalfa trials in which the plots are square of side 20 links, arranged in a 5×35 lattice with no discard strips. Strong positive temporal correlation is observed. The log likelihoods given in table 3 show strong spatial correlations in both years.

In the Rothamsted report 1925–26, yields of grain and straw are recorded for wheat on Sawyer field in 1925, followed by roots and tops for swedes in 1926. The same layout was used in both years, a regular 6×8 grid of plots of area 0.1 acre, with one corner plot not planted. For another corner plot, one of the weighings in 1925 is missing. Both corner plots are ignored in both years. All four temporal correlations are negative. The plan illustrated in the Rothamsted report shows square plots, i.e. 22 yards on each side, but the text does not confirm this. These are very large plots, so discard strips, if present, would be negligible by comparison. The analysis in table 3 assumes square plots with no discards.

5. Combination of information

For each of the examples described above plus four additional examples, table 3 gives the maximized marginal log likelihood for the four-parameter model (2.1). The maximized log likelihood is also given for the Bessel$_0$ sub-model with $\nu=0$, the exponential model with $\nu=0.5$, the Whittle model with $\nu=1$ and the conformal model with $\lambda=\nu=0$. The power model, in which the Matérn covariance function is replaced by $-|x-x'|^{2\nu}/\nu$, is also a sub-model of (2.1), provided that $\nu>0$, but this restriction is not imposed in likelihood calculations reported in table 3. The chief goal of this section is to examine whether the values in table 3 are consistent with the hypothesis of conformal invariance.

In comparing the various numbers, we should bear in mind that the 25 non-negative parameter pairs (λ_r, ν_r) for the column labelled Matérn are unrestricted, i.e. they may take different values for different crops, or different values for the same crop in different years or places. The columns labelled Bessel$_0$, exp'l and Whittle correspond to the sub-models $(\lambda_r, 0)$, $(\lambda_r, 1/2)$ and $(\lambda_r, 1)$, in which only the range depends on the crop. The column labelled 'power' corresponds to the model $(0, \nu_r)$, also with 25 unrestricted parameters. The column labelled 'conformal' is the intersection of Bessel$_0$ and power, in which $(0, 0)$ is the value for all crops in all seasons.

The first conclusion is that discrimination between these models is not easy, and that all are roughly equally effective as judged by the log-likelihood values
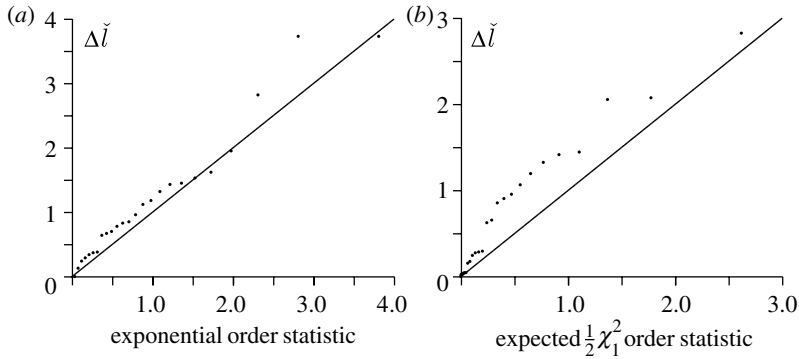
Figure 4. Ordered log-likelihood increments over the conformal model plotted against the nominal order statistic; (*a*) Matérn model and (*b*) power model.

achieved. Another way of saying the same thing is that the data appear to be at least approximately consistent with the hypothesis that $(\lambda_r, \nu_r) = (\lambda, \nu)$ is the same for all crops in all seasons, and in fact with the more specific hypothesis that $(\lambda, \nu) = (0, 0)$. To put this hypothesis to the test, we proceed as if the trials were independent. For the Matérn alternative, each likelihood ratio statistic would be distributed as $(1/2)\chi_2^2$ if this were a regular parametric problem. However, the conformal model is a boundary point in the space, and the range parameter may not be consistently estimable, so the nominal distribution may not be a satisfactory approximation. Nevertheless, this comparison is not an unreasonable one, so we use the standard graphical technique to compare the observed increments with the nominal distribution. The plot of the ordered log-likelihood increments against the exponential order statistics (figure 4*a*) shows no evidence of either systematic or individual departures from the nominal distribution. The two largest log-likelihood increments occur for the alfalfa trials, which are strongly correlated.

A similar comparison was made of the conformal model against the power model, for which the nominal distribution of increments is $(1/2)\chi_1^2$. Figure 4*b* shows the ordered log-likelihood increments plotted against the nominal expected order statistic. The maximum of 25 independent $(1/2)\chi_1^2$ variables has a mean of 2.62 and a standard deviation of 1.1, so the largest observed increment of 2.84 for the Wiebe wheat trial is not extreme relative to the nominal distribution. The Wiebe likelihood has its maximum at $(0, 0.19)$, or a power variance of 0.38. The second and third largest increments occur for the alfalfa trials. Ordinarily, in a half-normal plot, one tries to identify discrepant points, individual experiments for which the null model clearly fails, but figure 4*b* shows little evidence of individual departures of this sort. Instead, the evidence points to a small systematic departure in the central part of the distribution. The sum of the log-likelihood increments is 19.38, at the upper 3% point of the nominal $(1/2)\chi_{25}^2$ distribution. Temporal and other correlations have only a modest effect on the distribution of this statistic. The principal effect is to increase the variance additively by the sum of squared correlations, i.e. roughly from 12.5 to 16.0. The observed value of 19.38 is 1.72 standard deviations above expectation, suggestive but not strong evidence against the conformal hypothesis.
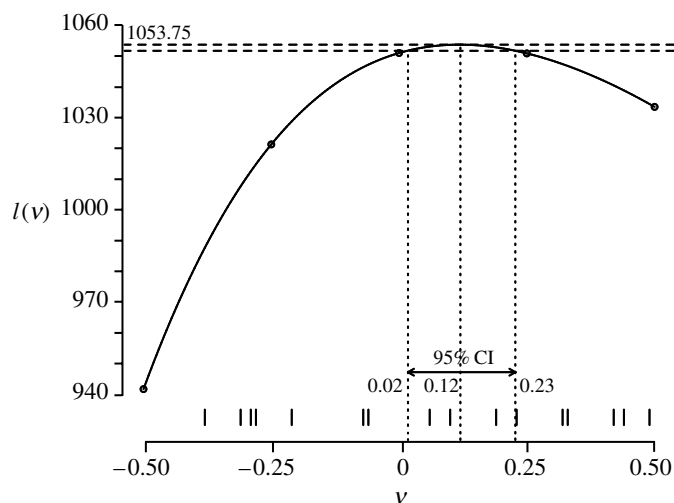
Figure 5. Profile log likelihood for ν in the power model using the combined data from all experiments. Vertical bars show individual estimates of ν in this range.

A small-scale simulation study for a 10×10 lattice of plots suggests that the null distribution of increments is roughly $0.4\chi_1^2$, somewhat smaller than the nominal distribution. By that yardstick, the observed total of 19.38 is in the extreme tail.

The preceding analysis suggests that the power model may be slightly superior to the conformal sub-model, either because $\nu_r = \nu$ is a non-zero constant or because the index varies from trial to trial. Accordingly, the power model with constant ν was fitted to all 25 trials as if they were independent, and the log likelihood is plotted against ν in figure 5. The maximum occurs at $\hat{\nu} = 0.12$, and the value $\check{l}(\hat{\nu}) = 1053.75$ is an increase of 2.7 log-likelihood units above the conformal model. The nominal 95% confidence interval for ν is (0.02, 0.23), i.e. powers in the range 0.04–0.46. The effect of correlations is to increase this interval slightly. Nine values of $\hat{\nu}_r$ are negative and 16 positive, the two smallest being $-0.64, -0.60$ and the two largest $0.96, 1.46$. Values in the range of the graph are plotted in figure 5.

To summarize, the information from the 25 trials is consistent with the sub-model in which both Matérn parameters are constant, the range being infinite, and the index parameter in the range zero to one-quarter. Although the exponential and Whittle models with variable range fit reasonably well, no exponential or Whittle model with constant λ is consistent with the observed data.

6. Anisotropy

So far, we have relied on the score test for geometric anisotropy as described in §2d. The values for the conformal model are shown in table 4.

For the mangold data, the values shown are for the conformal model with geometric anisotropy oriented along rows. For the three cases in which one or other component of the anisotropy statistic exceeds 1.96, the first component is

Table 4. Anisotropy test statistics for the conformal model.

Sp-B	1.12	1.33	MH-g	0.79	−1.43	Saw-g	−0.03	0.33
Ar-I	1.07	0.16	MH-s	2.14	−1.78	Saw-s	−0.17	0.40
Ar-II	−1.56	1.81	Man-r	1.22	0.92	Saw-t	−0.40	−0.02
Ant	2.40	0.46	Man-t	0.36	−0.61	Saw-r	−0.21	1.17
Eur	0.36	0.33	WW	−0.36	−0.43	Br-sp	0.63	1.81
Val	−1.17	0.70	Wb-w	−0.11	−1.18	Oat-g	−1.97	−0.33
Wal	0.47	−0.15	F-S-w	−0.61	−0.27	Oat-s	−0.69	−0.27
Jon	1.51	−1.76	Alf24	0.08	0.41	Pot	0.58	1.64
			Alf25	−1.85	−0.98			

the larger one. Thus, where anisotropy is present, it is associated with the orientation of rows and columns, and may indicate inaccuracies in the assumed width of discard strips.

For the barley data, a $6'$ discard strip between columns increases the log likelihood by 1.05 for the conformal model, and changes the test statistics to $(0.05, -1.78)$. For the Matérn and power models, the increase is 1.94. For a uniformity trial, a discard strip of this size is near the upper limit of the plausible range, and the computation gives an idea of the sensitivity of the test statistics to plot geometry. For the Antelope Heights navel orange data with plots $22' \times 22'$, an additional space of $5'$ between columns increases the log likelihood by 1.70 units to 28.61, and the test statistics are reduced to $(1.46, 0.48)$. The description given by Batchelor & Reed (1918) leaves no room for ambiguity concerning grid dimensions, so this appears to be a genuine anisotropy in the orchard, though not large in absolute terms. On the whole, the values reported in table 4 are little affected by modest variations in the assumed width of discard strips.

The likelihood ratio test shows moderately strong evidence of geometric anisotropy with row/column orientation in four cases with log-likelihood increments in the range 2.2–3.7, and strong evidence in two cases (M-H straw and F-S wheat) with increments of 6.6 and 6.3, respectively. Both halves of the Arlington orchard exhibit anisotropy, but the effects are in opposite directions. Discrepancies between the score test and the likelihood ratio test might be explained by the fact that the null hypothesis is not a regular sub-model.

In addition to geometric anisotropy, other sorts of anisotropies connected with rows and columns may also exist. One option is to use the non-isotropic $AR_1 \times AR_1$ model for spatial correlation, in which

$$\log K(x, x') = -\beta_1 |x_1 - x_1'| - \beta_2 |x_2 - x_2'|,$$

where x_1 is measured along rows and x_2 along columns (Cullis & Gleeson 1991). This is a four-parameter model with two variance components and two range parameters, all positive. The three-parameter conformal model with geometric anisotropy is not a sub-model, so log-likelihood differences are not to be treated as formal test statistics. In 17 of the 25 datasets, the difference in log likelihoods was less than one unit, and in four cases the conformal model had the larger value. Of the remaining eight cases, four had log-likelihood differences in excess of two units. These are MH-s, Man-r, Wiebe and Fairfield Smith's data, with log-likelihood differences of 2.5, 3.3, 27.6 and 3.0, respectively. Since the difference

was most striking for the Wiebe data, we decided to look more carefully at the source of this anisotropy.

One peculiarity of the Wiebe data is the plot shape, one foot in one direction with a single row of wheat, and 15′ in the other direction. Fairfield Smith's design is similar, in that each plot contains a single row of wheat, so anisotropy at this scale is inevitable. In an effort to accommodate this, we included in the covariance function of the conformal model indicator matrices for first-order neighbour plots in each direction. The coefficient for columns is negligible, but the coefficient for rows is large and positive, and the log likelihood is increased by 25.1 units. Aggregation of adjacent rows in sets of five reduces this effect substantially, but does not eliminate it. Thus, simple geometric anisotropy fails to accommodate the observed pattern of correlations at distances comparable to the drill width.

7. Estimation of effects

(*a*) *Example 22: brussels sprouts*

In each of the examples considered thus far, the plants have been grown under uniform conditions apart from possible row and column effects, which we have not attempted to model. Such experiments are ideal for studying spatial variation unencumbered by variety differences and other treatment effects applied to the plots. In this section, we examine briefly some of the issues connected with the estimation of variety and treatment effects in more typical field trials. One alternative to a spatial model is a randomized blocks model. Provided that the blocks are reasonably homogeneous and compact, the gain in precision from using a spatial model is unlikely to be very great in the majority of cases. Accordingly, gains from using a spatial model are likely to be greatest in a poorly designed experiment with badly chosen blocks, and we could demonstrate a large effect by choosing such a design. However, the examples we use are not of this type.

The Rothamsted Experimental Station Report (1934–36, p. 191) describes a randomized blocks design for brussels sprouts, using four blocks of 12 plots. Fortunately, the plot geometry is described in sufficient detail to permit spatial covariance models to be used. The plots are listed as 10×14 yds, but the actual harvested area is 9×13 yds after rejecting edge rows, so the plot separation is one yard. In addition to the control level, the eight treatments are, in order, sulphate of ammonia (1,2), poultry manure (3,4), soot (5,6), rape dust (7,8), with low and high levels of each. This is not a factorial design, so there are no treatment combinations. The response is the total weight in pounds of saleable sprouts for three pickings listed in table 5, with treatment level in parentheses. The design is unbalanced in the sense that there are four control plots in each block.

For examples of this sort, a model that eliminates additive row and column effects, as has been done in table 3, is unnecessarily wasteful. In each of the analyses in this section, row and column effects are treated as additive independent Gaussian variables. The model formula for the mean includes treatment effects only, and the model for the covariance matrix is

$$\sigma_0^2 |A| I_n + \sigma_1^2 |A|^2 V + \sigma_r^2 B_r + \sigma_c^2 B_c,$$

Table 5. Plan and yields in lbs of saleable sprouts.

235 (7)	213 (0)	216 (0)	273 (3)	268 (4)	243 (7)	240 (8)	243 (1)
214 (0)	238 (5)	237 (1)	243 (4)	240 (6)	255 (3)	204 (0)	220 (0)
237 (8)	245 (6)	237 (0)	251 (2)	255 (0)	250 (0)	260 (2)	246 (5)
257 (8)	258 (0)	254 (3)	266 (5)	263 (8)	234 (0)	240 (7)	253 (3)
266 (2)	266 (0)	256 (4)	254 (6)	259 (4)	249 (0)	265 (1)	227 (0)
260 (7)	248 (0)	228 (0)	235 (1)	248 (6)	270 (2)	250 (5)	236 (0)

a convex combination of four symmetric matrices in which $B_r(i, j) = 1$ if plots i, j are in the same row. With the response in kilograms and all lengths in metres, the estimated variance components are (0.105, 0.00131, 5.92, 1.05). The marginal log likelihood is 2.20 for the random blocks model, 4.06 for the conformal model and 5.30 for the extended version shown above. Alternative models implying dependence of successive row effects were also tried, but none of these led to substantial improvement of the fit. Block effects beyond the conformal model were found to be negligible.

(b) *Example 23–24: Long Hoos oats*

This example is an experiment conducted on Long Hoos field at Rothamsted in 1926 concerning the effects of early and late application of nitrogenous fertilizers on the yield of oats (Rothamsted Experimental Station Report 1925–26, p. 146). Data are available for both grain and straw on each of 96 plots. The field layout has 12 rows and eight columns, each plot one-fortieth of an acre (121 square yards) in area. For the actual design and original analysis, the field was partitioned into eight blocks (4×2), each block consisting of 12 plots. Since there are eight distinct treatment combinations, plus four controls labelled distinctly, a complete randomized blocks design was used.

Other than total area, the plot dimensions are not reported, nor is there any mention of discard strips. It appears that the practice was to use discard strips 3′–6′ wide, and our analysis assumes the smaller value. Under the assumption of isotropy in the conformal model, it is possible to infer the plot dimensions to a limited extent. This is done for both grain and straw, one providing a check on the other. The combined data suggest an aspect ratio of around 1 : 3, so the remaining analysis takes the plots to be 6×20 yards.

Unlike table 3, the analyses reported in table 6 take X to be the model matrix for treatments only. For grain yields, the marginal log likelihood is 11.70 for the random blocks model, 17.11 for the conformal model alone, 17.87 for *conformal+ block*, 17.17 for *conformal+row+col* and 18.20 for all terms combined (table 6). For straw, the marginal log likelihoods for the same models are 20.29, 48.80, 48.80, 49.26 and 49.26, respectively. Little evidence can be seen for row, column or block effects beyond natural isotropic spatial variation.

The average variance of treatment contrasts for the conformal model is 1.85 kg^2 for the conformal model, and 2.16 for the random blocks model. The values for straw are 7.4 and 15.4, respectively, so the estimated efficiency factor is 17% for grain and 100% for straw.

Table 6. Parameter estimates for two linear models.

| parameter | conformal + row + col | | random blocks | |
	estimate	s.e.	estimate	s.e.
control	0.00	—	0.00	—
sulphate 0	3.21	2.6	4.69	3.2
sulphate 1	9.72	2.7	12.28	3.2
manure 0	11.62	2.7	10.91	3.2
manure 1	7.82	2.8	9.89	3.2
soot 0	8.68	2.6	6.95	3.2
soot 1	4.14	2.8	5.47	3.2
rape dust 0	4.59	2.9	4.45	3.2
rape dust 1	6.24	2.9	6.61	3.2

(c) *Example 25: potatoes*

The final example is taken from a study by Eden & Fisher (1929) on the effects of nitrogen and potash on potato yields at Rothamsted in 1927. The layout of the plots in the field is 9×9, with nine blocks in a 3×3 arrangement. The plots are reported as being one-fortieth of an acre (121 square yards), but the actual plot dimensions are not given. The two principal treatment factors are the quantities of potash and nitrogen in cwt per acre, each with three levels. In addition, there are three kinds of potash salt, sulphate, muriate and 'low grade'. As usual, no distinction exists between zero dose of sulphate and zero dose of muriate, so the total number of treatment combinations is 21. In fact, the model matrix used here for treatment effects corresponds to the model formula *nitro.potash* + *kind*, so X has rank 11.

The 1927 Rothamsted Report gives little additional information regarding plot configuration or discard strips. The Rothamsted Experimental Station Report (1934–36) reports show that the standard drill width for potatoes was 3.5 links or $28''$. The harvested area typically consisted of the central $k-2$ drills in a plot containing k drills, so the discard strips are seven links wide between plots on the long side. We assume that there were similar discard strips on the short side. The reported plot area is sometimes given as the harvested area, and sometimes as the total area, but in the Eden–Fisher example, the one-fortieth acre clearly refers to the harvested region. Typically, k is in the range 5–8, and the aspect ratio of the harvested region ranges from $1 : 1.4$ (1936, p. 213), to $1 : 2.6$ (1935, p. 179), to $1 : 9$ (1934, p. 182). For the analysis shown here, the plots are taken to be 6×20 yards with discard strips seven links on all sides. The width of the plots must have been an integer number of drills, but this constraint is ignored in our analyses.

The marginal log likelihood is 11.82 for the random blocks model. 22.95 for the conformal model alone, 22.95 for *conformal* + *block*, 26.34 for *conformal* + *row* + *col* and 26.34 for all terms combined. The estimated variance component is 12.3 kg^2 for rows, and zero for columns and blocks. The coefficient $\hat{\sigma}_0^2$ is also zero, which is scarcely plausible. The average estimated variance of treatment contrasts in kilograms per plot is 35.3 kg^2 for the random blocks model, 20.4 kg^2 for the conformal model and 15.2 kg^2 for the model *conformal* + *row*. Evidently, blocking was not entirely effective in this instance.

8. Fairfield Smith's power law

According to Fairfield Smith (1938), the expected value of the sample variance of yields on n contiguous plots of area x behaves as $E(s^2) = c_n x^{1+b}$ for some index $0 < b < 1$, characteristic of the crop and season. In the investigation leading to this conclusion, a fixed total area T was partitioned into non-overlapping plots of area x by aggregation of smaller plots. The observed sample variance per unit area s^2/x^2 was plotted against area on a log–log scale, and for each of 40 trials, a straight line with slope in excess of -1 was obtained. Although $n = T/x$ is not constant in this study, the conclusion drawn was that $E(s^2) = c x^{1+b}$ for some index $b > 0$. For n contiguous plots of fixed area, the empirical evidence is very strong that the expected sample variance is an increasing function of n, so c_n is not constant. Likewise, although Fairfield Smith does not address this point, the index b in the power law may depend on T in addition to the crop and season.

According to the conformal model, the expected value of the sample variance of yields on an $m \times m$ array of contiguous plots of area x is

$$E(s^2) = \sigma_0^2 x + \sigma_1^2 x^2 m^2 \log(m)/(m^2 - 1),$$

where σ_0^2, σ_1^2 are constants independent of m and x. Under a model with kernel **1**, plot yields do not have a variance, so $E(s^2)$ is not to be confused with plot variance. For the Fairfield Smith investigation, $m^2 x = T$ is fixed and x/T is small, so the dependence on x is approximately

$$E(s^2/x) \propto 1 - \tfrac{1}{2}\eta(x/T)\log(x/T),$$

where $\eta = T\sigma_1^2/\sigma_0^2$ is a dimensionless constant proportional to the study area. A graph of $E(s^2/x)$ against x reveals an approximate power law over the range of interest $0 < x/T < 0.1$. For typical trials, the total area is such that $\eta > 1$, and the index in the power law is approximately a multiple of $\log \eta$.

Virtually, all of the processes described in this paper exhibit approximate power-law behaviour under plot aggregation. In the absence of treatment and other non-isotropic effects, the index is a simple function of model parameters. Without an adjustment for systematic effects, the graph of empirical variances against plot area is not an effective way of discriminating one process from another.

9. Summary and conclusions

Our analyses are based on the assumption that spatial variation of crop yields has a non-anthropogenic component that is isotropic and stationary as a planar process. Our hypothesis is that the non-anthropogenic component is also stationary as a spherical process. Additional sources of variation are invariably present, usually associated with the mechanics of planting and harvesting. The hope is that they can be identified and accommodated, in part, by the inclusion of additive row and column effects. This strategy is reasonably effective, but it does fail at scales comparable to the drill width. Anisotropy at larger scales of several drill widths does occur, but is usually less pronounced.

In almost all examples studied, the non-anthropogenic component of variation was found to be reasonably close to isotropic and reasonably close to conformally

invariant. In other words, regardless of the crop and the climate, spatial correlations at moderate to large distances decay logarithmically.

References

Andrews, D. F. & Herzberg, A. M. 1985 *Data*. New York: Springer.

Bartlett, M. S. 1978 Nearest neighbour models in the analysis of field experiments (with discussion). *J. R. Stat. Soc. B* **40**, 147–174.

Batchelor, L. D. & Reed, H. S. 1918 Relation of the variability of yields of fruit trees to the accuracy of field trials. *J. Agric. Res.* **12**, 245–283.

Besag, J. 1974 Spatial interaction and the statistical analysis of lattice systems (with discussion). *J. R. Stat. Soc. B* **36**, 192–236.

Besag, J. & Higdon, D. 1999 Bayesian analysis of field experiments using neighbouring plots (with discussion). *J. R. Stat. Soc. B* **61**, 691–746. (doi:10.1111/1467-9868.00201)

Besag, J. & Kempton, R. 1986 Statistical analysis of field experiments using neighbouring plots. *Biometrics* **42**, 231–251.

Besag, J. & Mondal, D. 2005 First order intrinsic autoregressions and the de Wijs process. *Biometrika* **92**, 909–920. (doi:10.1093/biomet/92.4.909)

Carathéodory, C. 1954 *Theory of functions*. New York: Chelsea.

Chilès, J.-P. & Delfiner, P. 1999 *Geostatistics*. New York: Wiley.

Clifford, D. 2005 Computation of spatial covariance matrices. *J. Comput. Graph. Stat.* **14**, 155–167. (doi:10.1198/106186005X27626)

Clifford, D., McBratney, A. B., Taylor, J. & Whelan, B. M. In press. Generalized analysis of spatial variation in yield monitor data. *J. Agric. Sci.*

Cressie, N. A. C. 1993 *Statistics for spatial data*. New York: Wiley.

Cruddas, A. M., Reid, N. & Cox, D. R. 1989 A time series illustration of approximate conditional likelihood. *Biometrika* **76**, 231–237.

Cullis, B. R. & Gleeson, A. C. 1991 Spatial analysis of field experiments—an extension to two dimensions. *Biometrics* **47**, 1449–1460.

de Wijs, H. J. 1951 Statistics of ore distribution. Part I: frequency distribution of assay values. *J. R. Neth. Geol. Min. Soc. New Ser.* **13**, 365–375.

de Wijs, H. J. 1953 Statistics of ore distribution. Part II: theory of binomial distribution applied to sampling and engineering problems. *J. R. Neth. Geol. Min. Soc. New Ser.* **15**, 125–124.

Diggle, P. 1988 An approach to the analysis of repeated measurements. *Biometrics* **44**, 959–971.

Eden, T. & Fisher, R. A. 1929 Studies in crop variation. *J. Agric. Sci.* **19**, 201–213. Reprinted in *Collected Papers of R. A. Fisher.* **2** (ed. J. H. Bennett) Adelaide University Press.

Fairfield Smith, H. 1938 An empirical law describing heterogeneity in the yield of agricultural crops. *J. Agric. Sci.* **28**, 1–23.

Gilmour, A. R., Cullis, B. R. & Verbyla, A. P. 1997 Accounting for natural and extraneous variation in the analysis of field experiments. *J. Agric. Biol. Environ. Stat.* **2**, 269–293.

Gleeson, A. C. & Cullis, B. R. 1987 Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. *Biometrics* **43**, 277–288.

Harville, D. A. 1977 Maximum-likelihood approaches to variance-component estimation and to related problems. *J. Am. Stat. Assoc.* **72**, 320–338.

Hille, E. 1965 *Analytic function theory*. New York: Blaisdell.

Jiang, J. 1996 REML estimation: asymptotic behavior and related topics. *Ann. Stat.* **24**, 255–286. (doi:10.1214/aos/1033066209)

Journel, A. G. & Huijbregts, C. J. 1978 *Mining geostatistics*. London: Academic Press.

Kempton, R. A. & Howes, C. W. 1981 The use of neighbouring plot values in the analysis of variety trials. *Appl. Stat.* **30**, 59–70.

Matérn, B. 1986 *Spatial variation*. New York: Springer.

McBratney, A. B. & Webster, R. 1981 Detection of a ridge and furrow pattern by spectral analysis of crop yield. *Int. Stat. Rev.* **49**, 45–52.

McCullagh, P. 2002 What is a statistical model? (with discussion). *Ann. Stat.* **30**, 1225–1310. (doi:10.1214/aos/1035844977)

Mercer, W. B. & Hall, A. D. 1911 The experimental error of field trials. *J. Agric. Sci.* **4**, 107–132.

Papadakis, J. S. 1937 Méthode statistique pour des expériences sur champ. *Bull. Inst. Amel. Plantes a Salonique* **23**, 13–29.

Patterson, H. D. & Thompson, R. 1971 Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554.

Pearce, S. C. 1978 Discussion of Bartlett (1978). *J. R. Stat. Soc. B* **40**, 159–160.

Rothamsted Experimental Station Report 1925–26. Harpenden: Lawes Agricultural Trust, p. 146.

Rothamsted Experimental Station Report 1934–36 Brussels sprouts: effect of sulphate of ammonia, poultry manure, soot and rape dust, pp. 191–192. Harpenden: Lawes Agricultural Trust.

Rudin, W. 1987 *Real and complex analysis* 3rd edn. New York: McGraw-Hill.

Stein, M. L. 1999 *Interpolation of spatial data*. New York: Springer.

Summerby, R. 1936 The value of preliminary uniformity trials in increasing the precision of field experiments. *MacDonald Coll. Tech. Bull.* **15**, 1–64.

Tørgersen, F. & Waagepetersen, R. 2004 Statistical modelling and deconvolution of yield meter data. *Scand. J. Stat.* **31**, 247–264. (doi:10.1111/j.1467-9469.2004.02-024.x)

Tunnicliffe Wilson, G. 1989 On the use of marginal likelihood in time series model estimation. *J. R. Stat. Soc. B* **51**, 15–27.

Wahba, G. 1990 *Spline models for observational data*. Philadelphia, PA: SIAM.

Whelan, B. M. & McBratney, A. B. 2002 A parametric transfer function for grain-flow within a conventional combine harvester. *Precis. Agric.* **3**, 123–134. (doi:10.1023/A:1013867418756)

Whittle, P. 1954 On stationary processes in the plane. *Biometrika* **49**, 434–449.

Whittle, P. 1956 On the variation of yield variance with plot size. *Biometrika* **43**, 337–343.

Whittle, P. 1962 Topographic correlation, power law covariance functions, and diffusion. *Biometrika* **49**, 305–314.

Wiebe, G. A. 1935 Variation and correlation in grain yield among 1500 wheat nursery plots. *J. Agric. Res.* **50**, 331–357.

Wilkinson, G. N., Eckert, S. R., Hancock, T. W. & Mayo, O. 1983 Nearest neighbour (NN) analysis of field experiments (with discussion). *J. R. Stat. Soc. B* **45**, 151–211.

Zimmerman, D. L. & Harville, D. A. 1991 A random field approach to the analysis of field-plot experiments and other spatial experiments. *Biometrics* **47**, 223–239.