Appl. Statist. (2002)
51, *Part* 1, *pp.* 23–36

The Bayesian choice of crop variety and fertilizer dose

Chris M. Theobald

University of Edinburgh, UK

and Mike Talbot

Biomathematics and Statistics Scotland, Edinburgh, UK

[Received April 2000. Revised June 2001]

Summary. Recent contributions to the theory of optimizing fertilizer doses in agricultural crop production have introduced Bayesian ideas to incorporate information on crop yield from several environments and on soil nutrients from a soil test, but they have not used a fully Bayesian formulation. We present such a formulation and demonstrate how the resulting Bayes decision procedure can be evaluated in practice by using Markov chain Monte Carlo methods. The approach incorporates expert knowledge of the crop and of regional and local soil conditions and allows a choice of crop variety as well as of fertilizer level. Alternative dose–response functions are expressed in terms of a common interpretable set of parameters to facilitate model comparisons and the specification of prior distributions. The approach is illustrated with a set of yield data from spring barley nitrogen–response trials and is found to be robust to changes in the dose–response function and the prior distribution for indigenous soil nitrogen.

Keywords: Bayesian inference; Decision theory; Hierarchical non-linear model; Local area prediction; Optimal fertilizer strategy

1. Introduction

1.1. Background

In agriculture, both economic and environmental considerations demand that chemical fertilizers should be applied only in quantities which are strictly justified. Even if environmental concerns are ignored, the use of some fertilizers, such as nitrogen, to increase crop yields shows diminishing and even negative returns at high dose levels, so the extra benefit from applying more fertilizer may be outweighed by the additional cost. The choice of how much fertilizer to use for a crop depends on the expected return at each dose level, on the cost of the fertilizer and on the amount of nutrient that is already in the soil.

To provide data on the expected gains from applying a fertilizer, trials are carried out using different dose levels, and these trials are often repeated over locations and/or years. The usual approach to determining the optimum level (as described, for example, in France and Thornley (1984)) is as follows:

(a) to specify a parametric dose–response model for the yield;
(b) to derive (as a function of the parameters) the dose level which maximizes the expected value of the crop minus the cost of applying fertilizer;

Address for correspondence: Chris M. Theobald, Department of Mathematics and Statistics, University of Edinburgh, James Clerk Maxwell Building, King's Buildings, Mayfield Road, Edinburgh, EH9 3JZ, UK.
E-mail: cmt@maths.ed.ac.uk

 0035–9254/02/51023

(c) to substitute parameter estimates based on the trial data into the expression for the optimum level.

Some reasons why this approach may be incomplete are as follows.

(i) The 'optimum' level determined in this way does not take account of sampling variation in the parameter estimates. We may investigate the likely cost of using estimates rather than true parameter values (as in Wallach and Loisel (1994)), but this does not necessarily lead to an improved choice of level.
(ii) To allow for (and to quantify) the effects of differences between locations, information from separate trials needs to be combined via a hierarchical dose–response model (Wallach, 1995a, b) rather than pooled as if the same parameter values applied to all environments.
(iii) There may also be information about the target environment which should influence the choice of fertilizer level, such as the result of a test of the amounts of nutrients already in the soil and crop yields in previous seasons. Babcock *et al.* (1996) derived a posterior distribution for soil nitrate based on data from a soil test and used this to estimate the optimum level by assuming a linear-plus-plateau dose–response function. Wallach (1995b) included a covariate, the clay fraction for each location, in a hierarchical model.

The present paper proposes a decision theory framework for fertilizer optimization which includes hierarchical modelling of trial data and the incorporation of information on soil nutrients in the target environment. It also addresses the following inadequacies of the conventional approach to this problem.

(iv) Information may be available on the responses of individual crop varieties. If the varieties have been grown in the same environments then it is efficient to analyse the data on all the varieties together.
(v) There may be a choice of the variety to be grown as well as of the fertilizer level to be applied.
(vi) Expert knowledge of the crop may indicate likely parameter values.

1.2. Application of Bayesian decision theory

Suppose that x_* and y_* denote the fertilizer level and the resulting yield (per unit area) for a particular variety grown in a target environment, and that the unit cost of the fertilizer and the value of a unit of the crop are denoted by c and p, so that the fertilizer cost and crop value for unit area are respectively cx_* and py_*. The values of c and p are known approximately when the choice of level is made. It would be straightforward to generalize the cost of fertilizer to allow for a non-linear dependence on x_* (such as a discontinuity at 0 representing the saving from not applying any fertilizer); other possible generalizations are to a vector of levels of different nutrients or to a single nutrient to be applied on different dates in the growing season.

Bayesian decision theory (as described for example in DeGroot (1970) or Cox and Hinkley (1974)) offers a solution to the determination of a genuine optimum fertilizer level (or combination of variety and level) given the trial data and the choices of parametric model and prior distribution. Under this prescription, we must do as follows:

(a) define the utility of applying level x_* at the target location: an obvious (but not automatic) choice is the crop value minus the cost of fertilizer for unit area, i.e.

$$u(x_*, y_*) = py_* - cx_*; \tag{1}$$

(b) specify a prior distribution for the model parameters which reflects knowledge of the dose–response relationship and of the target environment;
(c) derive the posterior distribution of the parameters given the set D of trial data;
(d) calculate the posterior expected utility $E\{u(x_*, y_*)|x_*, D\}$ at each of the possible fertilizer levels;
(e) choose the level giving the highest posterior expected utility.

If a choice is to be made of the variety to be sown as well as of the fertilizer level then the calculation in (d) must be carried out for each variety and the overall optimum combination found (taking account of any differences between varieties in the cost of seed and the price commanded for the crop).

If θ denotes the vector of parameters used in the model then calculating the posterior expected utilities is *not* achieved by replacing θ by an estimate, even a Bayesian estimate, in a formula for the expected utility given θ. In this respect our prescription extends the partly Bayesian approaches of Wallach (1995a, b) and Babcock *et al.* (1996) in which point estimates are required for some parameters.

Also the future crop yield would usually be assumed independent of the trial data D given θ. If this assumption is made, the posterior expected utility for equation (1) equals

$$p\, E\{E(y_* \mid x_*, \theta)|D\} - cx_*. \tag{2}$$

The inner expectation $E(y_* \mid x_*, \theta)$ in expression (2) is the dose–response function: in a hierarchical model for the trial and future yields, this would not depend on the parameters specific to other locations. The outer expectation is over the posterior distribution of θ. The *objective function* of Wallach and Loisel (1994) is $E(y_* \mid x_*, \theta) - cx_*/p$ in our notation.

Section 2 of this paper concerns the specification and fitting of dose–response models for past and future crop yields. Section 3 applies the ideas of Sections 1 and 2 by using data on the yield response of spring barley to nitrogen fertilizer and also examines the robustness of the resulting choice of level to changes in the dose–response function and the prior distribution for indigenous nitrogen. Section 4 discusses the desirability and feasibility of the Bayesian approach. The data which are analysed in this paper can be obtained from

`http://www.blackwellpublishers.co.uk/rss/`

2. Dose–response models for a single nutrient

We consider here dose–response models for a single nutrient and first distinguish a *dose–response function*, defining the possible dependence of expected yield on fertilizer level for a particular variety in an individual environment, from a *dose–response model*, which also describes other aspects of the relationships between the crop yields, including those between the parameters of the dose–response functions for different environments and varieties.

The following considerations are relevant to the choice of dose–response models under a Bayesian approach to fertilizer optimization.

(a) The dose–response functions should be sufficiently flexible to give a good representation of the dependence of crop yield on fertilizer level. For example, Boyd *et al.* (1976) found that strictly increasing functions, such as the exponential and inverse linear functions, tend to give a poor fit to data on spring barley. Quadratic functions (which are commonly used for fertilizer optimization) consistently overestimate the

fertilizer level for maximum yield: more realistic functions show a much more gradual change in expected yield above the turning-point than below it. Cerrato and Blackmer (1990) pointed out the invalidity of the quadratic model in the context of maize trials. Boyd *et al*. (1976) concluded that the model with two straight lines and the inverse quadratic function give the best fit. After examining a large set of trials, Sparrow (1979) recommended the inverse quadratic and the inverse linear functions with a descending linear function over the model with two straight lines.

(b) The models should be hierarchical to incorporate information from several environments, preferably relating to several locations and years. If effects for the environments are modelled as a random sample from a normal distribution, this variation should be on a suitable scale.

(c) Model parameters should be easily interpretable to simplify the specification of prior distributions.

2.1. Dose–response functions

We consider below three dose–response functions relating the expected yield to the dose x of a single fertilizer which allow a rapid increase to a maximum and a more gradual (or no) decrease beyond the turning-point. One, the linear-plus-plateau function, is specified by three parameters: the others require four. To facilitate the specification of prior distributions—and comparisons between the models—we express them all in terms of parameters which are intended to be easy to interpret. Pike (1984) has suggested a similar set of parameters for the inverse quadratic dose–response function.

The first of these parameters, representing the amount of indigenous nutrient per unit area, requires an assumption of additivity between the nutrient that is already in the soil and that which is added, so that it is meaningful to consider the *total* nutrient level after application: the definition might be easily generalized to allow for a conversion (possibly specific to the crop) between the values of indigenous and applied nutrient. We assume that the expected yield is 0 when the total nutrient level is 0, but positive above this value (except possibly at values well beyond the range of levels that are used in practice).

One advantage of this choice of parameter is that we may have information from a soil test on the indigenous nutrient level. As in Babcock *et al*. (1996), the posterior distribution from a measurement model of the level at the target environment can be used as the prior distribution in the dose–response model.

The first three parameters are as follows:

(a) α is the amount of the nutrient that is already in the soil, so the total nutrient level is $x + \alpha$;

(b) γ is the maximum expected yield, achieved at $x = \delta - \alpha$ if the function has a turning-point and otherwise at ∞;

(c) δ is the total nutrient level giving the maximum expected yield, for functions with a turning-point.

The choice of a fourth parameter—where this is required—is less obvious. For the inverse quadratic function, Pike (1984) suggested using as a parameter the slope β, say, at $-\alpha$, i.e. when the total nutrient level is 0. We have found that the specification of a genuine prior distribution for β is difficult because this part of the dose–response function is outside the experience of the expert. This parameter can also have a high (negative) posterior correlation with α. It may instead be easier to consider the likely reductions in yield from using total nutrient levels which

are close to δ, say 10 or 20 kg below δ. For dose–response functions which are approximately quadratic around the maximum, such as the inverse quadratic and inverse linear functions considered below, this reduction in yield is roughly proportional to the curvature of the function at the maximum. Thus for our fourth interpretable parameter we assume that

(d) ε is minus the second derivative of the dose–response function at its maximum.

We consider the following three dose–response functions, expressed in terms of the above parameters.

The *linear-plus-plateau*, three-parameter, curve comprises a linearly increasing portion and a constant portion, expressible as

$$E(y \mid x, \alpha, \gamma, \delta) = \begin{cases} \gamma(x+\alpha)/\delta & (0 < x \leqslant \delta - \alpha), \\ \gamma & (x \geqslant \delta - \alpha). \end{cases} \tag{3}$$

The *inverse quadratic* function, recommended by Nelder (1966), Boyd *et al.* (1976) and Sparrow (1979), is the ratio of a linear function of x and a quadratic function, and is expressible as

$$E(y \mid x, \alpha, \gamma, \delta, \varepsilon) = \frac{2\gamma^2(x+\alpha)}{2\gamma(x+\alpha) + \delta\varepsilon(x+\alpha-\delta)^2} \qquad (x > 0). \tag{4}$$

The *inverse linear with descending linear function*, suggested by Greenwood *et al.* (1971) and also recommended by Sparrow (1979), is the sum of an inverse linear function and a descending linear function, which may be combined as

$$E(y \mid x, \alpha, \gamma, \delta, \varepsilon) = \frac{\gamma(x+\alpha)\{2\gamma - \delta\varepsilon(x+\alpha-\delta)\}}{2\gamma(x+\alpha) - \delta^2\varepsilon(x+\alpha-\delta)} \qquad (x > 0). \tag{5}$$

Other dose–response functions are in use which are not so easily expressed in terms of the set of parameters defined above, e.g. an exponential function with an upper asymptote minus a multiple of x, which might be written as

$$E(y \mid x, \alpha, \zeta, \eta, \kappa) = \zeta[1 - \exp\{-\eta(x+\alpha)\}] - \kappa(x+\alpha) \qquad (x > 0).$$

Also, the above argument for using the curvature at the maximum as a parameter obviously does not apply to the function comprising two straight lines.

2.2. Specifying and fitting hierarchical models

We regard the vector of parameters, $[\alpha\ \gamma\ \delta]$ or $[\alpha\ \gamma\ \delta\ \varepsilon]$, for any of the above dose–response functions as specific to a particular environment, and we analyse the data recorded for the various environments by using a hierarchical random-effects model. Under such a model, effects for the trial and target environments are treated as arising from a common distribution. If there is a year × location structure in the data then factors for years and locations may instead be included in the random-effects model. If several varieties are considered then γ, δ and ε also depend on the variety.

We follow modern practice by representing our hierarchical models by using directed acyclic graphs and fitting the models by using Markov chain Monte Carlo (MCMC) methods. See Gilks *et al.* (1996) and Brooks (1998) for introductions to these methods. Fig. 1 shows such a graph for the linear-plus-plateau dose–response function given by equation (3) when yields y_{jl} have been achieved at fertilizer levels x_{jl} in environment j. The nodes of the graph (represented by circles) correspond to the experimental fertilizer levels and yield data,

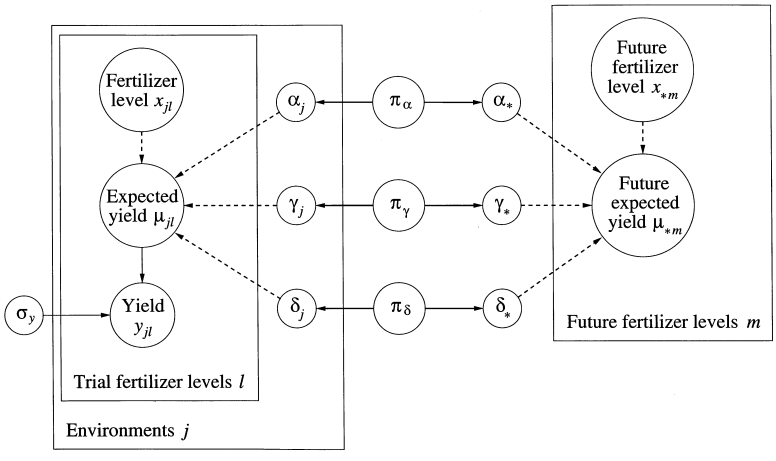

Fig. 1. Directed acyclic graph for applying model (3) in trial and target environments

the unknown parameters in the model and the possible future fertilizer levels and expected yields. Nodes within a rectangle are repeated over the levels of a factor, such as the environment classification. Full and broken arrows indicate stochastic and deterministic dependence respectively between nodes. Thus the distribution of y_{jl} is defined in terms of its expectation μ_{jl} and standard deviation σ_y, and the value of μ_{jl} is determined from x_{jl}, α_j, γ_j and δ_j. Higher level nodes π_α, π_γ and π_δ, each comprising one or more parameters, define the distributions from which the α_j, γ_j and δ_j are drawn.

The right-hand side of Fig. 1 includes 'future' nodes corresponding to the target location. In these nodes, index j for environments is replaced by an asterisk, whereas index l for experimental fertilizer levels is replaced by an index m defining the possible levels x_{*m} which might be used at this location. The value $(\alpha_*, \gamma_*, \delta_*)$ of (α, γ, δ) for the target location is assumed to be drawn from the same distribution as each of the $(\alpha_j, \gamma_j, \delta_j)$. To make predictions of future yields (rather than expected yields) we can add nodes y_{*m} with expectations μ_{*m} and standard deviation σ_y.

For the four-parameter dose–response functions (4) or (5), additional nodes for the ε_j and ε_* (and their parents) need to be included.

If several varieties are under consideration, we also require a variety factor in the graph with index i, say, and we seek to model recorded yields y_{ijl} corresponding to fertilizer levels x_{ijl} along with future expected yields μ_{i*m} and future yields y_{i*m}. In the above formulation, the parameters α_j relate only to the environments, but γ_j and δ_j (and ε_j if required) need to be generalized to γ_{ij} etc. to allow dependence on varieties, and the expected yields μ_{ijl} then depend on x_{ijl}, α_j, γ_{ij}, δ_{ij} and any ε_{ij}.

How the higher level nodes π_γ, π_δ and π_ε should be generalized when there is more than one variety depends on what we wish to assume about the effects of the varieties and their interaction with environments. We might, for example, assume that the γ_{ij} can be split into additive variety and environment effects γ_{vi} and γ_{ej} respectively, with a similar decomposition for the δ_{ij} (and the ε_{ij} if included in the model) so that

$$\left.\begin{aligned} \gamma_{ij} &= \gamma_{vi} + \gamma_{ej}, \\ \delta_{ij} &= \delta_{vi} + \delta_{ej}, \\ \varepsilon_{ij} &= \varepsilon_{vi} + \varepsilon_{ej}. \end{aligned}\right\} \tag{6}$$

Fig. 2 shows a possible expansion of the graph in Fig. 1 to include data on several varieties. It incorporates the additive model (6) for γ_{ij} and δ_{ij}, and makes the same prior assumptions about the performance of all the varieties, since each of the sequences of parameters $\{\gamma_{vi}\}$, $\{\gamma_{ej}\}$ etc. is assumed to be sampled from a common distribution. If the distributions of the α_j, γ_{vi}, γ_{ej} etc. are taken to be normal (conditional on their parent nodes) then their expectations and variances need to be specified. The common prior expectation of the γ_{ij} under the additivity assumption is the sum of those for the γ_{vi} and γ_{ej}, so only one of their prior expectations needs to be assumed non-zero: we take these expectations to equal a common value $\mu_{\gamma v}$ and 0 respectively. A similar argument applies to the δ_{ij} (and to any ε_{ij}).

Note that the nodes for the expected yields at the target location are connected directly to the variety components γ_{vi} and δ_{vi} but not to the environmental components γ_{ej} etc.: the variety to be grown in the target environment is one of those used in the trials, but the environment is assumed to be new.

Other elaborations of the graph may be useful, such as the following:

(a) including nodes for soil test results (as parents of the corresponding α-nodes);
(b) incorporating previous yields at the target location (as in Wallach (1995a)) by adding nodes for them or by including the target location as one of the trial locations;
(c) including nodes for future yields, so that their posterior predictive distributions can be estimated;
(d) allowing the utility to be non-linear in yield to reflect the farmer's aversion to risk (Anderson *et al.*, 1977);
(e) including dependence on covariate values for different environments;
(f) allowing the standard deviation of the yield to depend on one or more of the environment, the variety and the fertilizer level.

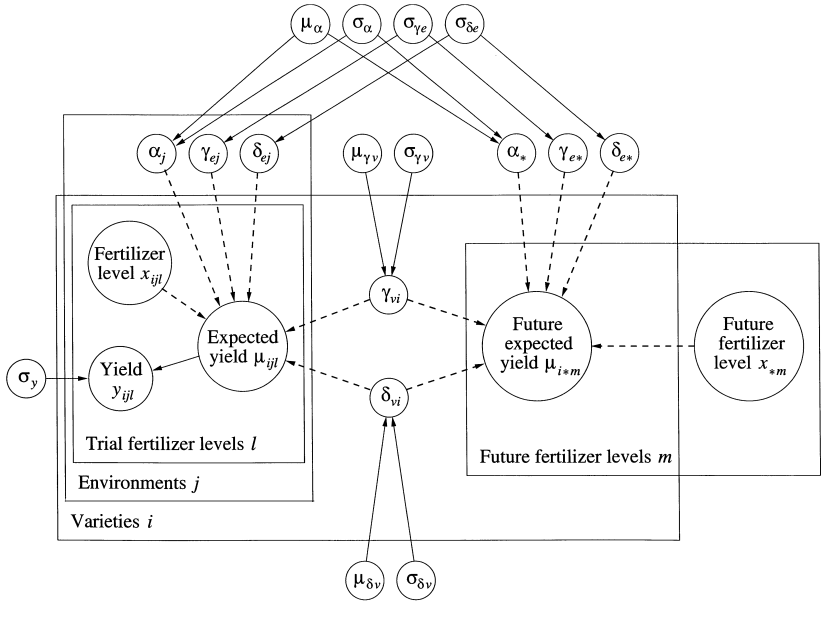

Fig. 2. Directed acyclic graph for applying model (3) with assumption (6) to several varieties in trial and target environments

We use the WinBUGS program (Spiegelhalter *et al.*, 1999) to fit our models to the yield data and to evaluate posterior distributions and expectations for any possible fertilizer level given the yield data. This program, which is freely available from

`http://www.mrc-bsu.cam.ac.uk/bugs`,

allows models to be specified by using a directed acyclic graph and selects and implements an appropriate MCMC method. It provides summary statistics and kernel density estimates of probability density functions for each node; the inclusion of nodes for the target location allows the calculation of posterior predictive distributions, posterior expected yields and hence posterior expected utilities for the possible fertilizer levels.

3. An example

Table 1 shows the grain yields (at 15% moisture content) of three varieties of spring barley (*Hordeum vulgare L.*), each grown with five levels of nitrogen fertilizer in two successive years at five trial centres in the Grampian region of north-east Scotland. Only seven of the 10 possible year × location combinations were used, but complete data were recorded for these. The yields show a substantial variation between the year × location combinations but suggest that variety Georgie tends to give the highest yields except at low nitrogen levels, where Sundance may be superior. However, Sundance and Midas are more suitable for malting than Georgie and may therefore command a higher market price.

Table 1. Grain yields of three spring barley varieties at five trial centres

Variety	*Centre*	*Yields (tonnes ha^{-1}) for the following applied nitrogen levels (kg ha^{-1}) and years:*									
		0		*35*		*62*		*90*		*125*	
		1976	*1977*	*1976*	*1977*	*1976*	*1977*	*1976*	*1977*	*1976*	*1977*
Georgie	1	4.206	5.710	4.586	6.014	5.075	6.609	5.795	5.527	6.009	4.249
	2	—	3.860	—	4.020	—	4.980	—	5.035	—	5.690
	3	4.785	—	5.159	—	5.791	—	5.681	—	5.594	—
	4	4.501	—	4.888	—	5.463	—	5.262	—	5.588	—
	5	5.625	3.164	5.897	4.097	6.785	4.769	6.823	5.113	7.219	5.365
Mean		4.55		4.95		5.64		5.61		5.67	
Midas	1	4.622	5.015	4.494	6.057	5.370	6.013	5.266	6.192	5.234	5.583
	2	—	3.370	—	4.165	—	4.835	—	5.485	—	6.125
	3	4.103	—	4.467	—	4.372	—	4.666	—	4.415	—
	4	4.217	—	5.079	—	4.884	—	4.928	—	5.099	—
	5	5.121	3.381	5.803	4.293	6.211	4.555	6.410	5.213	6.828	5.780
Mean		4.26		4.91		5.18		5.45		5.58	
Sundance	1	4.666	5.613	4.490	5.889	5.548	5.524	5.627	4.361	5.600	3.961
	2	—	3.690	—	4.170	—	4.790	—	5.180	—	5.595
	3	4.378	—	4.987	—	5.319	—	5.183	—	5.170	—
	4	4.369	—	4.872	—	5.197	—	5.118	—	5.535	—
	5	5.993	3.273	6.124	4.122	6.720	4.484	6.913	5.234	7.088	5.979
Mean		4.57		4.95		5.37		5.37		5.56	

3.1. Models and prior distributions

We evaluate and compare the posterior expected utilities corresponding to the three dose–response functions given by equations (3)–(5), carrying out the necessary calculations by using the WinBUGS program. We treat the year × location combinations merely as defining seven environments, since different fields were used in the two years.

As well as defining the nodes and their relationships via the directed acyclic graph, we must specify the distribution of each node given the values of any parents that it may have. The choice of these distributions allows expert knowledge of the crop to be incorporated: we eschew the use of improper or 'uninformative' prior distributions.

The distributional families that are used must be sufficiently flexible to be able to represent the prior opinions of experts, but sufficiently simple to allow these opinions to be expressed without undue difficulty. A convenient specification takes all the variances to have scaled inverse χ^2-distributions and the remaining stochastic nodes (including the trial and target yields) to be normal. Since our model is hierarchical, some unknown parameters have distributions defined by the values of their parents. We adopt the convention of using Roman symbols (with Roman or Greek subscripts) for the quantities which need to be specified by the expert to define the joint prior distribution for the model.

We have found that normal distributions are conveniently chosen by considering percentiles of the relevant distribution, whereas distributions for variances can be specified by giving a prior estimate of each variance and corresponding degrees of freedom. Thus, an estimate s^2 (with d degrees of freedom) of a variance σ^2 corresponds to $ds^2\sigma^{-2}$ having the distribution $\chi^2(d)$, so σ^{-2} has prior expectation s^{-2} and d measures the precision of this estimate. Equivalently, if $\mathrm{Ga}(a, b)$ denotes a gamma distribution with probability density function

$$b^a z^{a-1} \exp(-bz)/\Gamma(a) \qquad (z > 0)$$

then σ^{-2} has the distribution $\mathrm{Ga}(\frac{1}{2}d, \frac{1}{2}ds^2)$.

The distributions set out below in equations (7) correspond to the directed acyclic graph shown in Fig. 2 and hence assume a common prior distribution for the dose–response function of the three varieties, along with the additivity of variety and environment effects expressed in equations (6). The notation '$\sim$' is to be interpreted as 'is distributed as' or 'are distributed independently as', according to the context. Also the prior distributions for $\delta_{\mathrm{v}i}$, $\delta_{\mathrm{e}j}$, $\delta_{\mathrm{e}*}$ and their parents (and for $\varepsilon_{\mathrm{v}i}$, $\varepsilon_{\mathrm{e}j}$, $\varepsilon_{\mathrm{e}*}$ and theirs when relevant) are analogous to those given for $\gamma_{\mathrm{v}i}$, $\gamma_{\mathrm{e}j}$, $\gamma_{\mathrm{e}*}$, $\mu_{\gamma\mathrm{v}}$, $\sigma_{\gamma\mathrm{v}}$ and $\sigma_{\gamma\mathrm{e}}$. For each of the dose–response functions defined in equations (3)–(5), the expected trial yields μ_{ijl} are defined by replacing α, γ and δ (and ε, where required) by α_j, $\gamma_{\mathrm{v}i} + \gamma_{\mathrm{e}j}$, $\delta_{\mathrm{v}i} + \delta_{\mathrm{e}j}$ and $\varepsilon_{\mathrm{v}i} + \varepsilon_{\mathrm{e}j}$ respectively: for the target yields μ_{i*m}, they are replaced by α_*, $\gamma_{\mathrm{v}i} + \gamma_{\mathrm{e}*}$, $\delta_{\mathrm{v}i} + \delta_{\mathrm{e}*}$ and $\varepsilon_{\mathrm{v}i} + \varepsilon_{\mathrm{e}*}$.

$$\left.\begin{array}{lll} y_{ijl} \sim N(\mu_{ijl}, \sigma_y^2), & & d_y s_y^2 \sigma_y^{-2} \sim \chi^2(d_y), \\ \alpha_j, \alpha_* \sim N(\mu_\alpha, \sigma_\alpha^2), & \mu_\alpha \sim N(m_\alpha, t_\alpha^2), & d_\alpha s_\alpha^2 \sigma_\alpha^{-2} \sim \chi^2(d_\alpha), \\ \gamma_{\mathrm{e}j}, \gamma_{\mathrm{e}*} \sim N(0, \sigma_{\gamma\mathrm{e}}^2), & & d_{\gamma\mathrm{e}} s_{\gamma\mathrm{e}}^2 \sigma_{\gamma\mathrm{e}}^{-2} \sim \chi^2(d_{\gamma\mathrm{e}}), \\ \gamma_{\mathrm{v}i} \sim N(\mu_{\gamma\mathrm{v}}, \sigma_{\gamma\mathrm{v}}^2), & \mu_{\gamma\mathrm{v}} \sim N(m_{\gamma\mathrm{v}}, t_{\gamma\mathrm{v}}^2), & d_{\gamma\mathrm{v}} s_{\gamma\mathrm{v}}^2 \sigma_{\gamma\mathrm{v}}^{-2} \sim \chi^2(d_{\gamma\mathrm{v}}). \end{array}\right\} \tag{7}$$

The prior distribution used with all three dose–response functions is defined by specifying values for the quantities m_α, t_α^2, s_α^2 and d_α for the αs, $m_{\gamma\mathrm{v}}$, $t_{\gamma\mathrm{v}}^2$, $s_{\gamma\mathrm{e}}^2$, $d_{\gamma\mathrm{e}}$, $s_{\gamma\mathrm{v}}^2$ and $d_{\gamma\mathrm{v}}$ for the γs, analogous quantities for the δs and any εs, and s_y^2 and d_y for σ_y^2. Our values for these quantities, shown in Table 2, are based on experience with the crop in Scotland. In particular, the joint prior distribution for the $\varepsilon_{\mathrm{v}i}$ and $\varepsilon_{\mathrm{e}j}$ was chosen by considering the reduction in yield from using total nutrient levels that were 25 kg below δ.

Table 2. Parameter values defining the prior distribution for models (3)–(5) for spring barley

Parameter	*Mean*	*Variance*	*Estimate*	*Degrees of freedom*
μ_α	75	50	—	—
$\mu_{\gamma v}$	5.75	0.20	—	—
$\mu_{\delta v}$	150	625	—	—
$\mu_{\varepsilon v}$	3.2×10^{-4}	1.5×10^{-7}	—	—
σ^2_α	—	—	25	50
$\sigma^2_{\gamma e}$	—	—	0.25	20
$\sigma^2_{\gamma v}$	—	—	0.20	20
$\sigma^2_{\delta e}$	—	—	900	20
$\sigma^2_{\delta v}$	—	—	400	20
$\sigma^2_{\varepsilon e}$	—	—	2.0×10^{-8}	10
$\sigma^2_{\varepsilon v}$	—	—	5.0×10^{-9}	5
σ^2_y	—	—	0.25	100

A disadvantage of taking the prior conditional distributions of the α-, γ-, δ- and ε-parameters to be normal is that it allows them to take implausible or even nonsensical combinations of values in the posterior distribution, albeit with a very small probability. To prevent such values occurring in the calculation of the posterior expected utilities, we truncate the distributions of α_* and $\varepsilon_{vi} + \varepsilon_{e*}$ at zero and restrict the expected future yields μ_{i*m} to the interval [0, 20].

The models fitted make many assumptions, and we might assess the sensitivity of the posterior expected utilities to any of them. As well as fitting three dose–response functions, we consider the assumption that the common prior conditional distribution of the amounts of indigenous soil nitrogen α_j and α_* is normal. Babcock *et al.* (1996) suggested that a positively skewed distribution should be assumed for these levels and compared a uniform and a gamma prior distribution. Having assumed normality, we could investigate the sensitivity of our results to positive skewness in several ways: the method used here is to assume that the distributions in the second line of expression (7) apply instead to a monotone function ω, say, of α, so that the ω_j and ω_* have common prior distribution $N(\mu_\omega, \sigma^2_\omega)$, μ_ω has the distribution $N(m_\omega, t^2_\omega)$ and $d_\omega s^2_\omega \sigma^{-2}_\omega$ is distributed as $\chi^2(d_\omega)$. To attempt to isolate the effect of skewness from other differences between the distributions of α and ω, we must ensure that the two priors are similar in some sense, e.g. have the same first and second moments or the same values for specified quantiles. Here we illustrate taking ω to equal $\ln(\alpha)$ and match the two priors for α by taking d_ω equal to d_α and s^2_ω/t^2_ω equal to s^2_α/t^2_α, while keeping the 5% and 95% quantiles of the marginal distribution for α the same. The matching is achieved by simulating these distributions under the normal and log-normal models using WinBUGS.

For the values of m_α, t^2_α, s^2_α and d_α specified in Table 2, the distribution for $\ln(\alpha)$ which provides this match is given by $m_\omega = 4.298$, $t^2_\omega = 0.0136$, $s^2_\omega = 0.0068$ and $d_\omega = 50$.

3.2. Results

For each model, 100 000 MCMC iterates were used (with a Metropolis-within-Gibbs algorithm) following a 'burn-in' of 5000 iterates.

The posterior expected utilities are evaluated at applied nitrogen levels of 0, 10, …, 140 kg ha^{-1} and assume a price of £400 per tonne for fertilizer nitrogen: grain prices of £100 per tonne, £110 per tonne and £107.50 per tonne are assumed for Georgie, Midas and Sundance respectively, reflecting how suitable each variety is for malting. Figs 3(a)–3(c) show these posterior expected utilities under the three dose–response models with the parameter values specified in Table 2, along with the lower and upper 5% points of the posterior predictive distribution of utility. Corresponding plots assuming log-normal distributions for the α_j and α_* are almost identical with those assuming normality and are therefore omitted.

Because of their price advantage, the varieties Midas and Sundance achieve larger posterior expected utilities than Georgie, despite Georgie's generally higher mean yields. Table 3 gives the optimum levels (interpolated to the nearest kilogram per hectare) and the corresponding posterior expected utilities for each combination of the three varieties, the three dose–response functions and the normal and log-normal priors for α. Midas appears slightly superior to Sundance for all three functions and both priors, and would thus be selected for the optimum combination of variety and nitrogen level under the above price structure.

A comparison of results for the six combinations of dose–response function and distribution for α shows the following:

(a) given the remaining prior assumptions, the procedure is robust to replacing a normal prior for indigenous nitrogen by a matching log-normal prior;

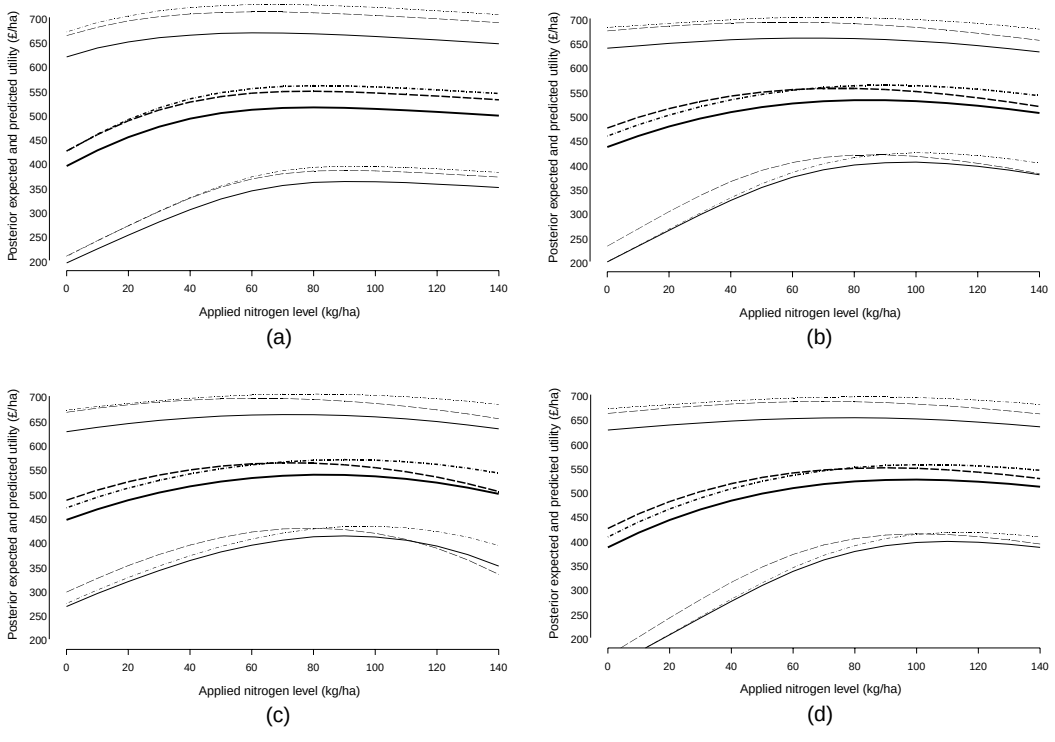

Fig. 3. Plots of posterior expected utility and the lower and upper 5% points of the posterior predictive distribution of utility for spring barley varieties Georgie (———), Midas (–·–·–·) and Sundance (– – –) using the dose–response functions (a) linear plus plateau, (b) inverse quadratic and (c) inverse linear with descending linear function: (d) shows the posterior expected utilities for these varieties under the inverse quadratic model when information on soil nitrogen at the target site is incorporated

Table 3. Optimum applied nitrogen levels and corresponding posterior expected utilities with normal and log-normal priors for indigenous nitrogen

Variety	*Prior for indigenous nitrogen*	*Optimum levels (kg ha^{-1}) and expected utilities (£ ha^{-1}) for the following dose–response functions:*		
		Linear plus plateau	*Inverse quadratic*	*Inverse linear*
Georgie	Normal	80, 515	85, 532	83, 538
	Log-normal	84, 519	85, 532	83, 538
Midas	Normal	82, 560	90, 563	91, 569
	Log-normal	87, 566	90, 562	91, 569
Sundance	Normal	79, 548	75, 556	73, 563
	Log-normal	83, 553	75, 556	73, 563

(b) very similar posterior expected utilities arise from the inverse quadratic dose–response function and the inverse linear with a descending linear function;
(c) the linear-plus-plateau function leads to smaller posterior expected utilities at low nitrogen levels than the other two functions;
(d) similar optimum levels are indicated under the three dose–response functions.

Observation (c) suggests that if the price of fertilizer nitrogen were much higher then the linear-plus-plateau function would indicate substantially lower optimum levels than the other two functions.

3.3. Using information on the target environment

These results assume that the environment in which the barley is to be grown is selected at random from the same population as the trial environments, but we may have available, and want to use, additional information on the amount α_* of indigenous nitrogen at the target environment. This might come from a soil test (as in Babcock *et al.* (1996)) or—less expensively—from knowledge of the site and of the crops grown in previous years. For example, if our assessment of α_* could be expressed as an estimate a_* with a standard deviation of t_a, we might take a_* to have distribution $N(\alpha_*, t_a^2)$ and add a node for a_* to Fig. 2 as a child of α_*. There are many alternatives to this assumption, of course: a_* might be given a Student or log-normal distribution, or taken to be normal with standard deviation proportional to its expectation, as in Babcock *et al.* (1996).

Fig. 3(d) shows the posterior expected utilities for the three varieties under the inverse quadratic model when the estimated level of indigenous nitrogen and its standard deviation equal 40 kg ha^{-1} and 5 kg ha^{-1}. The optimum levels and posterior expected utilities for Georgie, Midas and Sundance are respectively (99, 525), (103, 556) and (90, 550): Midas is still the best choice, but a higher applied nitrogen level is indicated because of the low level that is thought to be in the soil.

4. Discussion

The choice of fertilizer dose is one of many management issues in agriculture which is unambiguously a decision problem, rather than a problem of inference. The conventional approach to such problems involves deriving an optimum decision in terms of unknown

parameters and then replacing their values by point estimates. This leads to decisions which are optimum only in an asymptotic sense, which fail to exploit expert knowledge, and which may be based on unreasonable estimates of variance components in hierarchical models.

The Bayesian approach offers the possibility of incorporating expert knowledge of likely parameter values in a systematic way via the prior distribution. Although this prior information could be used to produce estimates with superior properties, even the use of improved estimates does not lead to an optimum decision: the Bayesian solution requires the integration of the utility function rather than the substitution of point estimates. Recent advances in Bayesian technology, particularly the development of MCMC methods, allow an increasing range of such decision problems to be tackled. The results obtained for the above example suggest that the resulting decisions may be robust to changes in the prior distribution and the dose–response model.

The incorporation of expert knowledge of how crop yields depend on indigenous and added nitrogen offers the possibility of reducing the size of fertilizer trials. For example, for most combinations of variety and environment the yields shown in Table 1 do not show a decline at the highest nitrogen level, so the data might be thought inadequate for a conventional analysis. We can use the sample data more efficiently by analysing all the yields in a hierarchical model and by adding information on the likely position and shape of the response curves.

The case for using prior distributions in decision-making is possibly stronger than in statistical inference (Cox and Hinkley (1974), page 417) but still requires careful modelling and elicitation of expert knowledge, possibly with co-operation between the subject-matter specialist and a statistician. Equally, the application of MCMC methods requires more judgment than the use of a conventional model fitting package.

Acknowledgements

The work was partially funded by the Home-Grown Cereals Authority and the Scottish Executive Environment and Rural Affairs Department. We thank Steve Hoad of the Scottish Agricultural College, Edinburgh, for useful discussions.

References

Anderson, J. R., Dillon, J. L. and Hardaker, J. B. (1977) *Agricultural Decision Analysis*. Ames: Iowa State University Press.

Babcock, B. A., Carriquiry, A. L. and Stern, H. S. (1996) Evaluation of soil test information in agricultural decision-making. *Appl. Statist.*, **45**, 447–461.

Boyd, D. A., Yuen, L. T. K. and Needham, P. (1976) Nitrogen requirements of cereals: 1, Response curves. *J. Agric. Sci. Camb.*, **87**, 149–162.

Brooks, S. P. (1998) Markov chain Monte Carlo method and its application. *Statistician*, **47**, 69–100.

Cerrato, M. E. and Blackmer, A. M. (1990) Comparison of models for describing corn yield response to nitrogen fertilizer. *Agron. J.*, **82**, 138–143.

Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. London: Chapman and Hall.

DeGroot, M. H. (1970) *Optimal Statistical Decisions*. New York: McGraw-Hill.

France, J. and Thornley, J. H. M. (1984) *Mathematical Models in Agriculture*. London: Butterworth.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996) *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

Greenwood, D. J., Wood, J. T., Cleaver, T. J. and Hunt, J. (1971) A theory for fertilizer response. *J. Agric. Sci. Camb.*, **77**, 511–523.

Nelder, J. A. (1966) Inverse polynomials, a useful group of multi-factor response functions. *Biometrics*, **22**, 128–141.

Pike, D. J. (1984) Discussion on a paper by Steinberg and Hunter. *Technometrics*, **26**, 105–109.

Sparrow, P. E. (1979) Nitrogen response curves of spring barley. *J. Agric. Sci. Camb.*, **92**, 307–317.
Spiegelhalter, D., Thomas, A. and Best, N. (1999) *WinBUGS User Manual*, *Version 1.2*. Cambridge: Medical Research Council Biostatistics Unit.
Wallach, D. (1995a) Empirical Bayes optimal fertilizer decisions. *J. Appl. Statist.*, **22**, 507–516.
———(1995b) Regional optimization of fertilization using a hierarchical linear model. *Biometrics*, **51**, 338–346.
Wallach, D. and Loisel, P. (1994) Effect of parameter estimation on fertilizer optimization. *Appl. Statist.*, **43**, 641–651.