HERITABILITY OF INTERNAL CHECKING IN *PINUS RADIATA*— EVIDENCE AND PRELIMINARY ESTIMATES

RODERICK D. BALL, MIKE McCONCHIE, and DAVE J. COWN

New Zealand Forest Research Institute,
Private Bag 3020, Rotorua, New Zealand

New Zealand Journal of Forestry Science 31(1): 78–87 (2001)

HERITABILITY OF INTERNAL CHECKING IN *PINUS RADIATA*— EVIDENCE AND PRELIMINARY ESTIMATES

RODERICK D. BALL, MIKE McCONCHIE, and DAVE J. COWN

New Zealand Forest Research Institute,
Private Bag 3020, Rotorua, New Zealand

(received for publication 18 December 2000; revision 31 May 2001)

ABSTRACT

Statistical techniques for assessing the evidence of heritability (h^2) of a highly non-normal random variable from limited-sized datasets were applied to internal checking in *Pinus radiata* D. Don wood. Bayesian hierarchical models for ordinal logistic regression with and without random family effects were fitted and compared using the technique of pseudo-priors to estimate the Bayes factor. Model parameters and Bayes factors were estimated using Gibbs sampling, and implemented using the computer program BUGS and importance sampling. Bayes factors of 203, 41.7 were obtained for the total number of checks and number of rings with checks, respectively. In the model assuming non-zero heritability, estimates and 95% credible intervals for h^2 were 0.64 (0.15–0.996) and 0.60 (0.06–0.997) respectively. In contrast, non-Bayesian methods including ANOVA with transformed variables and a non-Bayesian ordinal logistic regression analysis, failed to detect any effects for $p<0.05$.

Keywords: heritability; wood properties; Bayesian analysis; hierarchical models; BUGS; Gibbs sampling; pseudo-priors; ordinal logistic regression; non-normality; robust estimation; internal checking; oven-dry method; *Pinus radiata*.

INTRODUCTION

Recently there has been an increase in reported incidence of problems with internal checking in *Pinus radiata*. Researchers are attempting to understand the causes of checking so that the problem can be managed in current stands and avoided in future forests. In this paper we report our examination of statistical evidence for a genetic component of susceptibility to internal checking.

Internal checking, or *intra-ring checking*, is a wood quality problem whereby splits or checks occur within growth rings of wood. Internal checking affects the suitability of wood from pruned logs for high-value clearwood products from some stands. These checks may become visible only after machining or other expensive processing, so that even a small incidence of checking can have a significant quality cost.

The amount of checking, which usually occurs after kiln drying of wood (although checks in green timber have been observed from some of the worst-affected stands), depends on the severity of the drying schedule. A standardised treatment for oven drying of discs was developed by M.McConchie and reported by D.McConchie (1999). This treatment is quicker (cheaper) and also more severe than would normally be used, giving the opportunity to observe more checks of trees than would be found with typical wood processing, and hence to obtain more informative data on the susceptibility of trees to checking.

Two response variables were studied: the total number of checks, and the number of rings with checks. The total number of checks may be a good indicator of the quantity of pieces degraded when a log is cut into boards. The number of rings with checks is of interest because most checks take place in the sapwood just beyond the zone of heartwood formation. If the trees were allowed to grow on, these rings would form heartwood and most probably not form checks. Thus, trees with many rings with checks are considered more likely to have a checking problem at harvest than trees with many checks limited to a small number of rings.

Variance components require considerably more data to estimate than means. Therefore there is often considerable uncertainty associated with estimates of heritabilities, which are estimated as ratios of variance component expressions, yet heritability estimates are often given without standard errors. To study a non-normal random variable with a limited amount of data it is important to use a statistical method that does not rely on large sample sizes or normality for its validity, and can give distributional information on the estimate. The methodology which has these properties is Bayesian hierarchical modelling based on the ordinal logistic regression model with random effects with estimation using Gibbs sampling (*see*, e.g., Gelfand *et al.* 1990).

Bayesian analysis proceeds by combining prior knowledge (represented mathematically as a "prior distribution" on the unknown parameters) with the "likelihood" or statistical model which defines a probability distribution for the observed data in terms of the unknown parameters, obtaining a "posterior distribution" for the parameters which represents our knowledge *after* observing the data.

The output of the Gibbs sampler is a sample chain that is approximately a sample from the posterior distribution of model parameters. Estimates of the distribution of any quantity of interest can easily be obtained directly from this sample, avoiding the need to solve the integrals that are otherwise the main technical difficulty in applying Bayesian methods. Implementing the Gibbs sampler for a problem can be costly, requiring a significant amount of complex custom programming. If, however, a model can be specified and implemented using BUGS (a high level language and program for specifying and implementing a large class of Bayesian hierarchical models using Gibbs sampling — Spiegelhalter *et al.* 1995), results can be obtained with an order of magnitude less time and effort. Our preferred model will not run in BUGS, but fortunately a more generic model (with different prior distributions) will. The results for the preferred model can be derived from the BUGS output using importance sampling (*see*, e.g., Owen & Zhou 2000), which is a technique for sampling from one distribution, given samples from an approximating distribution.

Many papers have been written estimating heritability. Most modern approaches use REML (Patterson & Thompson 1971) to estimate the variance components. Recently, some authors have used Gibbs sampling to estimate genetic parameters, e.g., Magnabosco *et al.* (2000) and Sorensen *et al.* (1995).

A general outline of the method and the results are presented here; further details, including the implementation of the Gibbs sampler in BUGS and the Bayes factor calculation have been given by Ball (in prep.).

MATERIALS AND METHODS

Material was obtained from a "268 series" factorial trial of *P. radiata* growing on a site in the central North Island of New Zealand. The trial consisted of single full-sib families in rectangular blocks of $7 \times 7 = 49$ trees. The families assessed had been selected, from within a geographically relatively uniform area of the trial, to give a range of wood properties for a study of juvenile wood. The families were not unrelated: there were 23 distinct parents with some parents involved in several families, compared with the 36 distinct parents which would be needed for unrelated families. Eight trees per family were selected from the inner 5×5 array within each block.

Discs for internal check assessment were collected from the breast height position of butt logs on the eight trees (all 14 years old in 1999) from each of the 18 selected families (144 trees). Discs were split in half, and one half was assessed for internal checking using the "oven dry method" developed by M.McConchie and described by D.McConchie (1999). The number of checks was recorded for each annual ring of each disc; however, only the total number of checks per disc and the number of rings with checks are analysed here.

Statistical Methods

The number of checks can be very large in some trees, giving a highly non-normal distribution with a high degree of within-family variation in some families. To avoid problems with distributional assumptions not being satisfied due to the somewhat wild nature of the distribution, the number of checks and number of rings with checks were analysed as categorical variables. For the total number of checks, the categories used were 0, 1–5, 6–20, and > 20 checks per disc. For the number of rings with checks, the categories used were 0, 1–3, and > 3. These category sizes correspond roughly to none, low, medium, or high levels of checking.

The categorical data were analysed by ordinal logistic regression (McCullagh 1980), where category probabilities are related through the logistic link function to a linear model or "linear predictor". The linear predictor is similar to the linear mixed model which would be fitted in a REML analysis of normal data. Two models were fitted: Model 1 with no genetic variance (or $h^2 = 0$) and random within-family random error effects with variance σ_e^2, and Model 2 where, additionally, a random family effect for each family (with variance σ_f^2) was fitted.

The heritability was estimated as

$$h^2 = \frac{2\sigma_f^2}{\sigma_f^2 + \sigma_e^2} \quad (1)$$

A Bayesian approach was taken. To assess the evidence for non-zero heritability, the Bayes factor

$$B_{21} = \frac{\Pr(\text{data} \mid \text{Model 2})}{\Pr(\text{data} \mid \text{Model 1})} \quad (2)$$

was estimated. The numerator of (2) is the probability of the data under Model 2, and the denominator is the probability of the data under Model 1. Thus, the Bayes factor measures how much more likely the data are under Model 2 than Model 1.

The other main quantity of interest is the posterior distribution of h^2 in Model 2.

In order to apply Bayesian statistics it is first necessary to specify prior distributions for parameters in the model, whose distributions are not determined by other parameters. Following the recommendations of Jeffreys (1961), the Jeffreys non-informative prior for a proportion is used for h^2, and diffuse priors (containing little or no information) are given for the parameters θ_j (cut-points in the ordinal logistic regression) and the within-family variance σ_e^2. This gives a standard or "objective" approach to estimating the Bayes factor which can be applied in the absence of prior information.

Posterior distributions of parameters for each model, which represent our knowledge about a parameter after using the data, were estimated by Gibbs sampling and importance sampling. The method of pseudo-priors (Carlin & Chib 1995) was used to construct a Gibbs sampler for estimating the Bayes factor (*see* Ball in prep. for details).

RESULTS

Counts of the number of trees in each internal checking class for each family are given in Table 1.

Boxplots (Hoaglin *et al.* 1983) of the total number of checks assessed for each tree are shown in Fig. 1, and for the number of rings with checks in Fig. 2. The solid dots represent the medians of the data and the limits of the boxes are given by the upper and lower quartiles. The dotted lines or whiskers give an indication of the range of the data and are obtained by

TABLE 1–Number of trees per family in each internal checking class

Family	Total checks				No. rings with checks		
	0	1–5	6–20	>20	0	1–3	> 3
1	3	1	2	2	3	3	2
2	3	1	2	2	3	3	2
3	3	2	1	2	3	3	2
4	3	0	4	1	3	2	3
5	6	2	0	0	6	2	0
6	6	1	1	0	6	2	0
7	7	1	0	0	7	1	0
8	5	2	1	0	5	3	0
9	2	1	4	1	2	5	1
10	5	2	1	0	5	2	1
11	6	2	0	0	6	2	0
12	7	0	0	1	7	1	0
13	5	1	1	1	5	1	2
14	4	1	2	1	4	2	2
15	4	2	1	1	4	4	0
16	1	0	3	4	1	3	4
17	4	1	2	1	4	2	2
18	7	1	0	0	7	1	0

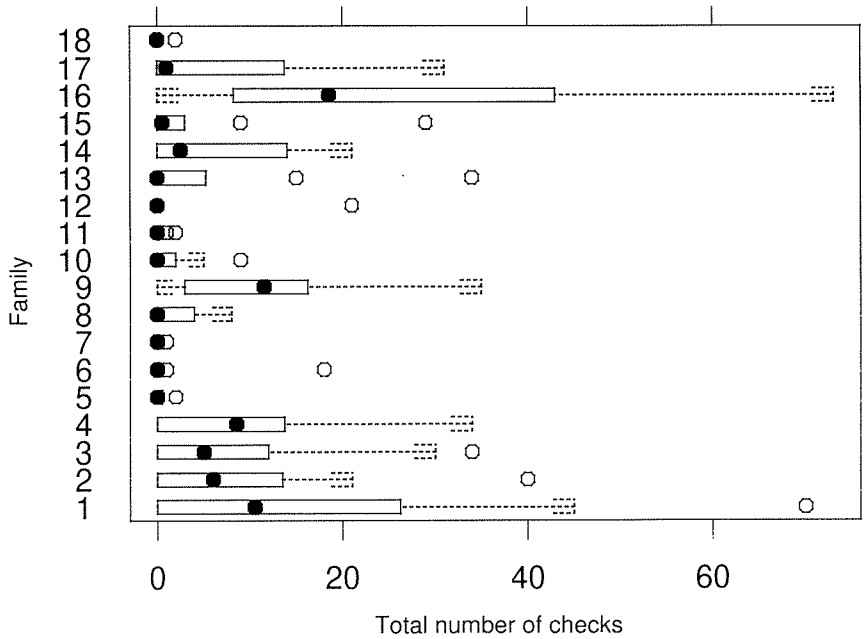

FIG. 1–Boxplots of total number of internal checks per tree by family.

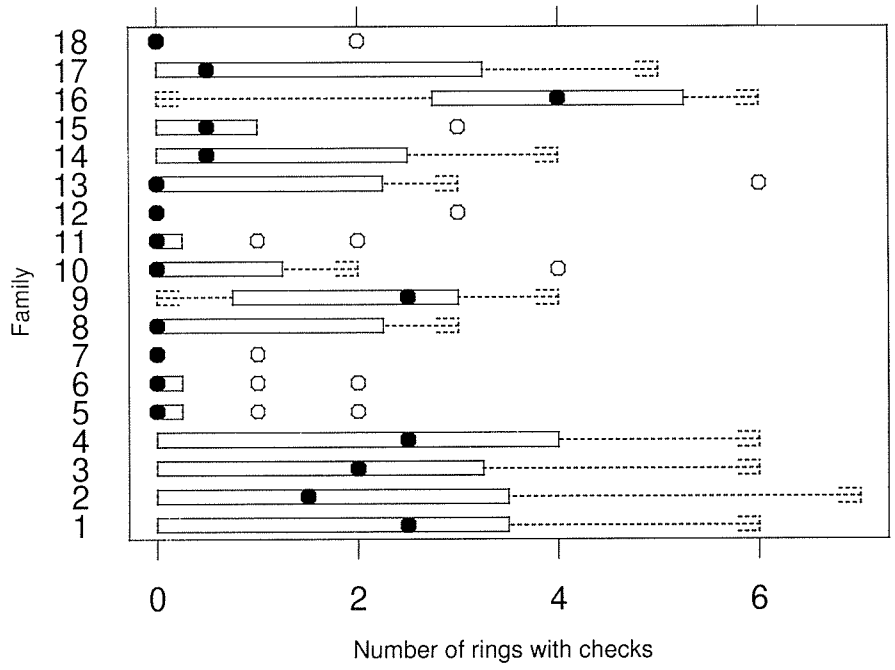

FIG. 2–Boxplots of the number of rings with checks per tree by family.

extending 1.5 times the interquartile range from the quartiles, then moving back to the nearest data point. Data points (possible outliers) outside the whiskers are plotted individually with an open circle.

Note the wide range in the total number of checks per tree within families. All families had one tree or more with checks and one tree or more with no checks. The best two families had only one tree with checks, while the worst family had seven out of eight trees with checks (Table 1). Family 12 had only one tree with checks, but this tree was in the worst checking class.

For Models 1 and 2, from 400 000 iterations of the Gibbs sampler the Raftery and Lewis convergence diagnostic (Raftery & Lewis 1992) indicated that the median could be estimated to within 5% and the 2.5%, 97.5% quantiles to within 1.25% for each model parameter in under 200 000 iterations.

For the Bayes factor estimation runs, from 1 000 000 iterations the Raftery and Lewis convergence diagnostic indicated that the proportion of times Model 1 was selected was estimated to within ±0.023, 95% of the time.

Parameter Estimates

Posterior model statistics for h^2 are shown in Table 2.

The prior and posterior distributions for h^2 for the total number of checks and number of rings with checks are shown in Fig. 3. It is instructive to compare the prior and posterior distributions for a parameter, to see the influence of the information from the data.

Posterior probabilities for Model 2 and Bayes factor estimates are given in Table 3.

TABLE 2–Posterior statistics for h^2. Estimates are based on 400 000 iterations of the Gibbs sampler for Model 2.

Total number of checks			Number of rings with checks		
Mean	95% ci	Median	Mean	95% ci	Median
0.64	(0.153–0.996)	0.66	0.60	(0.062–0.997)	0.60

DISCUSSION AND CONCLUSIONS

Some families had only one tree with a low number of checks and the rest with no check, suggesting family differences (Fig. 1) to many observers. However, the nature of variability of internal checking makes these differences difficult to pin down statistically. The "standard" statistical methods which were tried, including a naïve ANOVA with the original data transformed and a non-Bayesian ordinal logistic regression analysis, failed to detect any effects for $p<0.05$.

In general, categories were chosen to give a reasonable number of observations in each class. As with transformations, a different choice is effectively a different trait which would result in different statistics being obtained. Thus, it is desirable to standardise the category definitions.

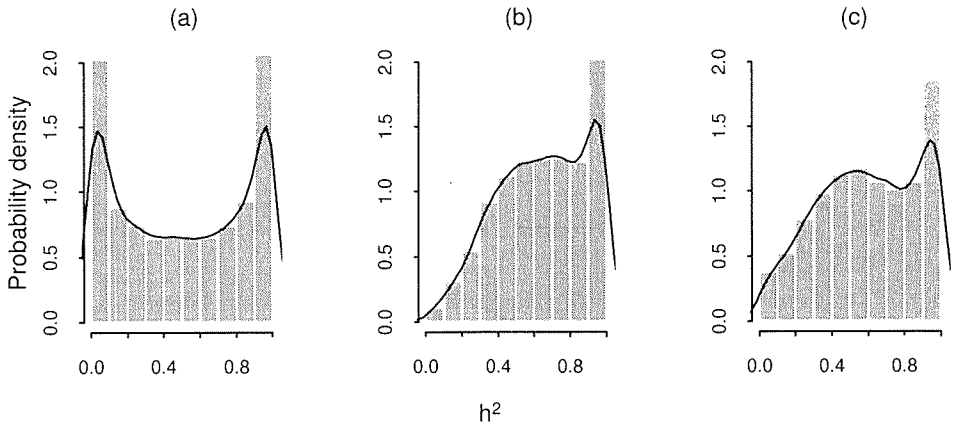

FIG. 3–Prior and posterior distributions h^2. The prior distribution for h^2 is shown in panel (a), and posterior distributions for h^2 for the total number of checks and number of rings with checks are shown in panels (b), (c).

TABLE 3–Posterior probability for Model 2 ($h^2 > 0$) and Bayes factor estimates. Estimates are based on 1 000 000 iterations of the Gibbs sampler for comparing Models 1 and 2, using the method of pseudo-priors. Posterior probabilities assume equal prior probability of 0.5 for Models 1 and 2.

	$Pr(h^2 > 0)$	95% ci	Bayes factor	95% ci
Total number of checks	0.995	(0.994–0.996)	204	(183–228)
Number of rings with checks	0.976	(0.974–0.978)	41	(38–44)

The method of this paper shows evidence for non-zero heritability of internal checking. Based on a non-informative prior for the within-family variance and a non-informative Jeffreys prior for h^2, the Bayes factors of 204, 41, were obtained representing strong evidence for heritability for both traits.

Implicit in the model is the genetic prior constraint $0 \leq h^2 \leq 1$. This genetic prior constraint certainly holds for the true heritability, although the proportion of variance that is within family could conceivably be higher than 1/2 if, for example, there were maternal effects. These are not considered likely, but if they do exist would not be estimable in the current experiment. Further investigation showed Bayes factors with the constraint are almost twice the unconstrained values—demonstrating that it is certainly worth using prior information, if available.

The heritability of internal checking may be quite high. The heritability of the linear predictor for the total number of checks under Model 2 was estimated to be 0.64 with a 95% credible interval of 0.15–0.996, and 0.60 with a 95% credible interval of 0.06–0.997 for the number of rings with checks. (Note: unlike confidence intervals from ANOVA, the credible intervals automatically lie within the range of the parameter, i.e., are between 0 and 1.)

Some wood-quality traits, including internal checking, are substantially more expensive to assess than the traits commonly assessed in tree breeding programmes. This study was not intended to be definitive, but rather to provide preliminary evidence for heritability to support the case for further studies, using material available from a study designed for other purposes. Consequently, the experimental design was not determined for the purpose of estimating heritability of internal checks and the sample sizes were not as large as we would like. The experimental design consisted of unreplicated single family blocks, which is not good for estimating heritability, as block level spatial variation is confounded with between-family variation. This may result in an upwards bias in our heritability estimates, and evidence for heritability is, strictly speaking, evidence for either heritability or block-level spatial variation. It can be argued that if spatial variation has been appropriately modelled, the remaining variation is genetic. We are currently attempting to fit 3–6 "super-blocks" which would strengthen the case that the block-level variation is likely to be family variation. The best approach is, however, to obtain further data from more suitable trials.

The expense of wood quality assessments also means that the number of trees assessed per family is likely to be too low to rely on the Central Limit Theorem to give approximate normality. Consequently, an approach such as that used here will be necessary.

There were 23 distinct parents with parents involved in several families, compared with the 36 distinct parents which would be needed to have unrelated parents. We have not considered the inter-relationship between families here: this could be incorporated in the model but only a small difference (increase) in the heritability estimate would be expected. Our model is based on conditional independence—conditional on the parental genotypes, the progeny in a family are independent. This still holds, and within family variance is unaffected, when some parents are common between families. The between family variance is estimated in our model from the variation between family means, which are the mean of parental values. The sample spaces for a genetic sample of family means obtainable from 23 parents ($1/2 \times 23 \times 22 = 253$ possible matings) and that obtainable from 36 parents ($1/2 \times 36 \times 35 = 630$ possible matings) are both close to representative of the infinite population.

Our interest is in establishing evidence for between-family variation rather than in a precise estimation of heritability for the purpose of estimating breeding values or genetic gain. Since we have full-sib families, our heritability will contain some non-additive variation in addition to the additive variation and hence would be expected to be intermediate between the narrow-sense heritability which determines the gain from family selection and broad-sense heritability which applies when selecting clones. To obtain unbiased estimates of the narrow-sense heritability, half-sib families are needed; however, twice as many half-sib families would be needed to obtain equivalent precision.

Although we have evidence for heritability, the precision of our estimates of heritability is low, as indicated by the range of the 95% credible intervals in Table 2, and further investigation with larger sample sizes is warranted. Comparable data from a sample size of 72 families (four times the size of the present study) would give heritability estimates with standard errors approximately half the size of those in Table 2, or approximately 0.13, comparable to most heritability estimates. The main proviso is that the incidence of checking must be at a level to enable efficient estimation of checking. The data analysed here, with some checking expressed in all families, and at least one tree from each family without checking, would be close to optimal for these purposes.

The Gibbs sampler for the ordinal logistic model with random family effects was successfully implemented in BUGS and enables estimation of posterior distribution of heritability of internal checks. There was no problem fitting the model to binary multinomial data derived from the raw trait values, but some prior information or bound on the variance components is needed to ensure parameters remain in a numerically feasible region.

Sorensen *et al.* (1995) developed a Gibbs sampler to estimate sire effects and variance components for a "threshold model" with categorical data, assuming the observed categories result from a "liability" (an unobserved normally distributed normal random variable) which was treated as an unknown parameter for each observation in their model. Their underlying model was, effectively, similar to our ordinal logistic regression model. By comparison, our approach is more direct, avoiding the need for unknown "liability" parameters in the model. They reported "inferences in good agreement with approximate maximum likelihood" for their example, with a total of 2674 offspring in 82 half-sib families. In general, for parameter estimates and confidence intervals, this is expected with sufficiently large sample sizes.

Bayesian and non-Bayesian methods for testing "precise hypotheses" such as no-heritability *vs* heritability are substantially different, however. To assess evidence for heritability, the Bayes factor comparing the probability of the data under alternative models is estimated. The Bayes factor gives a more rigorous and meaningful test for scientific hypotheses than traditional hypothesis tests and p-values (Berger & Berry 1988). For example, the interpretation of the Bayes factor as evidence does not depend on sample size; that for p-values does.

Previously, non-Bayesian methods had the advantage in computability and range of methods available. Now, with modern computers and the availability of algorithms and software for Gibbs sampling, and other Monte Carlo Markov chain methods, the computational difficulties of Bayesian modelling can be overcome, using conceptually simple but computationally intensive methodology. Bayesian methods now have the advantage in terms of ability to fit realistic models, and ability to compute probabilities without the need for asymptotic and distributional assumptions. This is in addition to the advantage of posterior probability distributions for statistical inference or optimal decision-making that they have always had. We recommend them to the readers of this journal.

ACKNOWLEDGMENTS

The authors would like to thank Dr Rowland Burdon, Dr Satish Kumar, and Dr Phillip Wilcox for useful comments which have improved the manuscript.

REFERENCES

BALL, R.D.: Assessing evidence for heritability of non-normally distributed traits (in prep.)

BERGER, J.; BERRY, D. 1988: Statistical analysis and the illusion of objectivity. *American Scientist 76*: 159–165.

CARLIN, B.P.; CHIB, S. 1995: Bayesian model choice via Markov chain Monte Carlo Methods, *Journal of the Royal Statistical Society B57(3)*: 473–484.

GELFAND, A.E.; HILLS, S.E.; RACINE-POON A.; SMITH, A.F.M. 1990: Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association 85*: 972–985.

HOAGLIN, D.C.; MOSTELLER, F.; TUKEY, J.W. (Ed.) 1983: "Understanding Robust and Exploratory Data Analysis". Wiley, New York.

JEFFREYS, H. 1961: "Theory of Probability", 3rd ed. Oxford University Press, London.

MAGNABOSCO, C.D.U.; LÔBO, R.; FAMULA, T. 2000: Bayesian inference for genetic parameter estimation on growth traits for Nelore cattle in Brazil, using Gibbs sampler. *Journal of Animal Breeding & Genetics 117*: 169–188.

McCONCHIE, D.L. 1999: An update on within-ring internal checking including a procedure to identify logs prone to form checks during drying. Proceedings of 3rd Wood Quality Symposium, "Emerging Technologies for Evaluating Wood Quality for Wood Processing". New Zealand Forest Research Institute, Rotorua.

McCULLAGH, P. 1980: Regression models for ordinal data. *Journal of the Royal Statistical Society B 42*: 109-142.

OWEN, A.; ZHOU, Y. 2000: Safe and effective importance sampling. *Journal of the American Statistical Association 95*: 135–143.

PATTERSON, H.D.; THOMPSON, R. 1971: Recovery of interblock information when block sizes are unequal. *Biometrika 58*: 545–554.

RAFTERY, A.E.; LEWIS, S.M. 1992: How many iterations in the Gibbs sampler? Pp. 763–773 *in* Bernardo, J.M.; Berger, J.O.; Dawid, A.P.; Smith, A.F.M. (Ed.) "Bayesian Statistics", Vol. 4. Oxford University Press, United Kingdom.

SORENSEN, D.A.; ANDERSEN, S.; GIANOLA, D.; KORSGAARD, K. 1995: Bayesian inference in threshold models using Gibbs sampling. *Genetics Selection and Evolution 27*: 229–249.

SPIEGELHALTER, D.; THOMAS, A.; BEST, N.; GILKS, W. 1995: BUGS—Bayesian inference using Gibbs sampling Version 0.50. Technical Report, MRC Biostatistics Unit, Cambridge. http://www.mrc-bsu.cam.ac.uk/bugs.