the geneticssociety

www.nature.com/hdy

ARTICLE

Evaluation of alternative methods for estimating the precision of REML-based estimates of variance components and heritability

Roshan Rameez[1], Shrinivas Jahageerdar[1✉], Jayasankar Jayaraman[2], Thongam Ibemcha Chanu[3], Rama Bangera[4] and Arthur Gilmour[5]

© The Author(s), under exclusive licence to The Genetics Society 2022

Residual Maximum Likelihood (REML) analysis is the most widely used method to estimate variance components and heritability. This method is based on large sample theory under the assumption that the parameter estimates are asymptotically multivariate normally distributed with covariance matrix given by the inverse of the information matrix. Hence, these sampling variances could be biased if the assumption of asymptotic approximation is incorrect, especially when the sample size is small. Though it is difficult to assess the impact of sample size, an alternative option is to generate a full distribution of the parameters to determine the uncertainty of estimates. In this study, we compared the REML estimates of variance components, heritability and sampling variances of body-weight (BW), body-depth (BD), and condition-factor (K) with those obtained from four sampling-based methods viz., parametric and nonparametric bootstrap, asymptotic sampling and Bayesian estimation. The aim was to understand if a sample size of order 1413 was sufficient to contain adequate information for a reliable asymptotic approximation. The REML solution was close to values obtained from different sampling-based methods indicating that the present sample size was sufficient to estimate reliable genetic variation in different traits with varying heritability. The variance and heritability estimated by a nonparametric bootstrap estimate based on randomization of family effects gave comparable results as evaluated by REML for different traits. Hence, the nonparametric bootstrap estimate can be effectively used to understand whether the sample size is large enough to contain sufficient information under likelihood estimation assumptions.

Heredity (2022) 128:197–208; https://doi.org/10.1038/s41437-022-00509-1

INTRODUCTION

Variance components and their functions are essential parameters of interest in selection experiments. In recent decades linear mixed model (LMM) analysis has become a popular tool for analyzing breeding data. The LMM analysis involves estimating variance components and predicting breeding values assuming the estimated variance components are correct. Several methods are available for the estimation of variance components (Searle et al. 2009), but Residual/Restricted Maximum Likelihood (REML) is the method of choice in pedigreed selection experiments (Lynch and Walsh 1998). The REML is a variation of the Maximum Likelihood (ML) method with an edge over the ML method in that the REML estimates of variance parameters are unbiased by the estimation of fixed effects (Patterson and Thompson 1971). The caveat is that the significant properties of the likelihood estimators are valid when the sample size approaches infinity, and the behavior of the same when working with small sample size is primarily unknown (Psutka and Psutka 2019). The likelihood theory suggests that the parameter estimates asymptotically have a multivariate normal distribution with covariance matrix given by the inverse of information matrix, i.e., the inverse of the matrix of second partial derivatives of the likelihood function (Meyer and Houle 2013). An estimate of sampling variance, a measure of the reliability of the parameter estimates (variance components and their functions), can be obtained from the inverse of the information matrix.

Though the estimation of sampling variance is straightforward, it is difficult to calculate reliable confidence intervals around functions (heritability) of these parameters (Waldmann and Ericsson 2006). Heritability, defined as the ratio of additive genetic variance to total phenotypic variance, is an important genetic parameter used to assess the potential for genetic improvement. The standard procedure to obtain the sampling variance of heritability is to linearly approximate the function with its first-order Taylor series expansion and then calculate the variance of this linear approximation (Meyer and Houle 2013) by the Delta method (Lynch and Walsh 1998). The REML analyses implemented in standard software packages such as WOMBAT (Meyer 2007), DMU (Madsen et al. 2014), ASREML (Gilmour et al. 2015), Echidna (Gilmour 2018), among others use the Delta method. However, the sampling variance might be biased under an incorrect asymptotic approximation, especially when sample size is small (Thai et al. 2013).

Moreover, while there are large-sample approximations for the sample variance of a REML variance estimator, it is unclear as to what amount of data constitutes a large enough sample size

[1]ICAR-CIFE, Mumbai, India. [2]ICAR-CMFRI, Kochi, India. [3]ICAR-CIFE, Kakinada, Andhra Pradesh, India. [4]BENCHMARK GENETICS, Bergen, Norway. [5]Statistical and ASReml Consultant, Keiraville, NSW, Australia. Associate editor Yuan-Ming. ✉email: jahageerdar@cife.edu.in

Received: 24 February 2021 Revised: 28 January 2022 Accepted: 2 February 2022
Published online: 23 February 2022

(Walsh and Lynch 2018). There are no direct approaches available to estimate the optimum sample size for a meaningful asymptotic approximation of likelihood estimates; hence, the large sample size is often a gray area. The multivariate normality of the likelihood estimates (the assumption under which the sampling variance of parameter estimates are quantified) is only valid when the sample size approaches infinity. Often, there is a lack of sufficient data to fulfill the conditions of optimal likelihood estimates. Therefore, the use of alternative approaches to evaluating uncertainties in variance components and their functions is desirable. Sampling-based methods are an option for estimating uncertainties associated with parameter estimates. Sampling approaches are usually aimed at generating a full distribution of parameter estimates instead of the point estimators. The full distribution of estimates can provide a complete picture of the uncertainties associated with estimating unknown parameters.

We aimed to compare the sampling variance in the estimates of variance components and heritability for three different traits (body weight with high heritability, body depth with medium heritability, and condition factor with low heritability) in *Clarias magur* (an Indian catfish species) obtained from the REML method with those obtained from sampling-based methods to assess whether the present moderate sample size was large enough to obtain a reliable asymptotic approximation. It was hypothesized that if the sample size contains sufficient information, the estimates obtained from various approaches should agree. Four well-known sampling-based methods were compared with the REML estimates viz., nonparametric bootstrapping, parametric bootstrapping, asymptotic sampling and Bayesian method to obtain a comprehensive picture of the uncertainty in the estimates of variance components and heritability.

MATERIALS AND METHODS

Dataset

The data used in the present study were from the base population of *Clarias magur* from an ongoing genetic selection program under ICAR-Central Institute of Fisheries Education, Mumbai, India. The Indian catfish, *C. magur*, is an economically important species with a high aquaculture potential. The fish used in the present study consisted of 78 fullsib families produced across 2 years (2014 and 2015) following a single pair mating design (each male mated with only one female—supplementary material attached as Annexure-1). The brooders were collected from three natural populations from three distinct geographical locations, viz., Andhra Pradesh, Assam, and West Bengal. Of the total 2328 fingerlings stocked, the measurements from 1413 survived fish were made after a year of communal rearing under mono and polyculture systems in different earthen ponds. The details of breeding, larval rearing, PIT tagging and communal rearing used in this study are described elsewhere (Jousy et al. 2018; Rameez et al. 2020). This study focused on the three important traits in aquaculture, body weight (BW), body depth (BD), and condition factor (K).

The animal model

An animal model in matrix notation takes the following form,

$$y = X\beta + Zu + e$$

where y is a vector of observations on all individuals, β is a vector of fixed effects (stock, batch, pond, sex, and body-weight at stocking as linear covariate), X represents a design matrix (principally 0s and 1s) relating the fixed effects to each individual, u is a vector of random animal effects, Z is a design matrix relating the random effects to each individual and e is a vector of residual errors. It is assumed that random effects u and e are normally distributed with mean zero and variance-covariance matrix G and R respectively, $u \sim N(0, G)$ and $e \sim N(0, R)$.

G is a matrix that is made of additive genetic covariances between relatives i and j, which is given by twice the coefficient of coancestry ($2\Theta_{ij}$ − the probability that an allele is drawn at random from individual i will be identical by descent to an allele drawn at random from individual j) times the additive genetic variance in the base population, i.e., $G = 2\Theta_{ij}\sigma_A^2$ or $G = A_{ij}\sigma_A^2$, where A is known as the additive genetic (or numerator) relationship matrix and has elements $A_{ij} = 2\Theta_{ij}$. e is assumed to be independently and identically distributed (iid). Hence the residual variance-covariance matrix is given by $R = \sigma_e^2 I$, where σ_e^2 is the residual variance and I is the identity matrix.

REML estimation of variance components

Mathematically REML is based on the linear transformation of the observation vector y to y^* that removes the fixed effects from the model, with the help of a transformation matrix K such that $KX = 0$ (Lynch and Walsh 1998). Applying the transformation matrix then yields the model equation as $y^* = KZu + Ke$ and the REML estimates are essentially the ML estimates for these transformed variables (Kruuk 2004). The transformation matrix K is also called the matrix of error contrasts, and hence REML is sometimes known as Residual Maximum Likelihood (Littell et al. 2006). The variance components using the REML method were estimated in software Wombat (Meyer 2007) using an animal model. For an animal model with a single random effect, the REML analysis will provide estimates σ_A^2 and σ_e^2, from which the heritability was estimated as $h^2 = \sigma_A^2/(\sigma_A^2 + \sigma_e^2)$.

Prediction of random effects

In an animal model, the random effects of interest are known as the breeding values of individual animals. An individual's breeding value for a given phenotypic trait is the total additive effect of its genes on that trait (Falconer and Mackay 1996). A general method for predicting random effects, known as the Best Linear Unbiased Predictor (BLUP), predicts individuals' breeding values from field records of large and complex pedigrees (Lynch and Walsh 1998). Robinson (1991) lists four different derivations for BLUPs of which the one developed by Henderson (1973) can obtain BLUP for random effects simultaneously with the Best Linear Unbiased Estimators of the fixed effects and hence the method is popularly known as Henderson's mixed model equation (MME). The MME for a general mixed model setting is:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\beta} \\ \widehat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

Suppose we assume that residual variance is iid, in that case, the R matrix can be factored out, and the MME for the animal model will reduce to (Isik et al. 2017):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\alpha \end{bmatrix} \begin{bmatrix} \widehat{\beta} \\ \widehat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

where, $\alpha = \frac{\sigma_e^2}{\sigma_A^2}$ or $(1 - h^2)/h^2$. The solutions to these equations are the best linear unbiased estimators of $\widehat{\beta}$ and the best linear unbiased predictors of $\widehat{u}$ for a given value of α. The BLUP breeding values were predicted by solving MMEs in Wombat (Meyer 2007) using estimated REML variances.

The standard error of heritability

Heritability is a nonlinear function (ratio) of variance components. The usual way to estimate the variance of a ratio is by the Delta method (Lynch and Walsh 1998). The Delta method uses the first-order Taylor series expansion of the function about the estimated $\widehat{\Theta}$ parameter (Meyer and Houle 2013). For the nonlinear function heritability, $h^2 = f(\Theta) = \sigma_A^2/(\sigma_A^2 + \sigma_e^2)$, provided $f(\Theta)$ is differentiable and continuous around the estimated Θ, the variance of heritability is given as (Stefan 2017):

$$Var(h^2) \approx \left(\frac{\partial h^2}{\partial \sigma_A^2}\right)^2 \sigma_{\sigma_A^2}^2 + \left(\frac{\partial h^2}{\partial \sigma_e^2}\right)^2 \sigma_{\sigma_e^2}^2 + 2\left(\frac{\partial h^2}{\partial \sigma_A^2}\right)\left(\frac{\partial h^2}{\partial \sigma_e^2}\right) \sigma_{\sigma_A^2,\sigma_e^2}$$

the square root of which gives the asymptotic standard error of heritability.

The functions, $\left(\frac{\partial h^2}{\partial \sigma_A^2}\right) = \frac{\sigma_e^2}{(\sigma_A^2+\sigma_e^2)^2}$ and $\left(\frac{\partial h^2}{\partial \sigma_e^2}\right) = \frac{-\sigma_A^2}{(\sigma_A^2+\sigma_e^2)^2}$ are the partial derivatives of the function of heritability with respect to the estimated additive genetic variance and error variance, respectively. The notations, $\sigma_{\sigma_A^2}^2$, $\sigma_{\sigma_e^2}^2$ and $\sigma_{\sigma_A^2,\sigma_e^2}$ are the variance of additive genetic variance, the variance of the error variance and covariance between additive genetic variance and error variance, respectively, all of which can be obtained from the asymptotic variance-covariance matrix. Based on the standard errors

from the delta method, ~95% confidence intervals (CI) were computed as $h^2 \pm 1.96\widehat{\sigma}_{h^2}$.

Nonparametric bootstrap (NPBS). A Nonparametric subject-wise bootstrap, wherein samples were drawn from the predicted random animal effects (BLUPs) and residuals, was performed. The resampling unit was a parent and its offspring, taking into account the family structure. The NPBS was performed under three different schemes. A total of 10,000 new datasets were generated, and for each data set, heritability was estimated using REML to obtain the sampling distribution of heritability estimates. The bootstrap samples were obtained as follows:

1. Fit the model, $y = X\beta + Zu + e$, to the data
2. Predict the random animal effects $\widehat{u}$ and residuals $\widehat{e}$
3. Obtain vector of fixed effects as $\widehat{y} = y - (\widehat{u} + \widehat{e})$, where $\widehat{y}$ is y vector corrected for effects of $\widehat{u}$ and $\widehat{e}$
4. Draw random animal effects with replacement from predicted $\widehat{u}$ (step 2) to obtain new $\widehat{u}^*$
5. Draw residuals with replacement from predicted $\widehat{e}$ (step 2) to obtain new $\widehat{e}^*$
6. Create new y values $\widehat{y}^*$ as $\widehat{y}^* = \widehat{y} + \widehat{u}^* + \widehat{e}^*$
7. Fit the model $\widehat{y}^* = X\beta + Zu + e$ to the new data set to estimate variance components and obtain heritability
8. Repeat the steps from 4 to 7 for 10,000 times

The NPBS was performed in six ways viz., two methods of analysis (animal model and family model) by three sampling schemes:

(1) u and e treated independently; bootstrap samples built separately from u and e without randomizing the families
(2) Linkage between u and e maintained; bootstrap samples built from $u + e$ without randomizing families and
(3) Linkage between u and e maintained; bootstrap samples built from $u + e$ randomizing the families.

The separation of the animal effect into genetic (u) and residual (e) components via the mixed model equations is in fact, just a computational strategy for fitting an animal effect that is correlated within families. In this data, where the genetic relationships are essentially defined as fullsib families (4 of the 78 are half-sib), the family model and animal model are (almost) equivalent, and the essential information required for bootstrapping is given by $u + e$ nested in families.

Parametric bootstrap. Here, new data sets were generated by sampling values for the animal effects and residuals for the data and pedigree structure from a multivariate normal distribution with a mean zero and covariance matrix as estimated from the REML analysis, a form of model-based simulation. The parametric bootstrap sample was obtained as follows:

1. Fit the model $y = X\beta + Zu + e$ to the data
2. Estimate the overall mean $E(y) = X\beta$
3. Generate random animal effect $\widehat{u}$ and residuals $\widehat{e}$ from the multivariate normal distribution
4. Create new $\widehat{y}$ values as $\widehat{y} = E(y) + \widehat{u} + \widehat{e}$
5. Fit the model $\widehat{y} = X\beta + Zu + e$ to the new data set to estimate variance components and obtain heritability
6. Repeat steps 3 to 5 for 10,000 times

Asymptotic sampling

A total of 10,000 sets of the variance components (both additive genetic variance and residual variances) were sampled from the multivariate normal distribution. The MVN was parameterized with the mean and covariance matrix obtained from the REML analysis. The heritabilities were estimated from the samples of variance components obtained in Wombat (Meyer 2007).

Bayesian estimation

The *MCMCglmm* package in R software, a combination of Gibbs sampling, slice sampling and Metropolis-Hastings updates, was used to generate the chain (Hadfield et al. 2019). The starting values generated by the default heuristic techniques within *MCMCglmm* were used to initialize the chain. A single chain was generated with a chain length of 10,00,000 iterations with a thinning interval of 100 iterations (which means only one iteration value is saved for every 100 iterations) to reduce the autocorrelation between successive values (de Villemereuil 2012). A burn-in period of 50,000 iterations was used to ensure the convergence of the chain before saving iteration values. A total of 9500 values of variance estimates were sampled from the chain for every trait, and heritability was estimated for each sampled variances.

An inverse-gamma distribution was used as the prior for variance components, and a diffuse normal prior centered around zero with a very large variance (10^8) was used for the mean (Hadfield et al. 2019). The inverse-gamma distribution was parameterized by two parameters nu (shape) and V (scale), where $\mathrm{nu} = 0.002$ and $V = 1$. The inverse-gamma distribution allows for a weakly informative prior on variance components and U-shaped prior (with a very steep shape on the borders in 0 and 1) on the heritability (de Villemereuil 2012). For algebraic convenience, under the Bayesian method, the trait values of K was first multiplied with a constant (100 in this case) before estimation of the variance components.

Convergence was assessed visually by looking at trace plots of parameters. The convergence diagnostic as proposed by Geweke (1992) for Markov chains was used to statistically test the chain for convergence. It was based on a test for equality of the means of the first and last part of a Markov chain (by default, the first 10% and the last 50%). Geweke's statistic was estimated as a statistical test for convergence using *gewke.diag()* function from the *coda* package in R. A thinning interval of 100 was used to reduce the autocorrelation and get a better effective sample size, wherein only one iteration value was picked from every 100 iterations. The thinning also reduces the memory required to hold results and lightens downstream analysis.

A 95% Highest Density Region (HDR) is interpreted as a $1 - \alpha$ (0.95) probability that the interval contains the actual value of the unknown parameter as opposed to the 95% CI from the frequentist method where $1 - \alpha$ (0.95) of the time the confidence intervals will enclose the parameter. HDRs are estimated by the *HPDinterval()* function from the *coda* package in R. The mode of the posterior distribution was estimated using the function *posterior.mode()* from the *coda* package in R. The kernel density plots were obtained using *proc kde* in SAS 9.3.

RESULTS

REML estimates

The results from the REML analysis of variance components and heritability are presented in Table 1. The additive genetic variance (255.20) and the residual variance (325.20) for the body weight and other traits (Table 1) were obtained by REML using the average information algorithm, and the corresponding sampling variances were obtained from the inverse of the average information matrix. The heritability estimates of 0.44 for BW, 0.22 for BD, and 0.08 for K were obtained as the function of variance components (ratio of additive genetic variance and total phenotypic variance), and their sampling distributions were obtained by approximating the function with its first-order Taylor series expansion. The results and their corresponding approximate 95% CI are given in Table 1.

MCMC diagnostics

Convergence and autocorrelation. The additive genetic variances, residual variances and heritabilities for body weight are presented in Fig. 1, plotted against iteration number. This shows the fluctuation in sampled values over the iterations (50,000 to 1,000,000 after dropping the first 50,000 iterations, known as burn-in). The trace plot depicts a well-mixing chain, which moves through the entire subset of parameter space without settling to any particular region. The MCMC trace plots for BD and K also followed the same pattern (not presented). The Geweke's test statistic was a standard Z-score; all the p values for the Z scores from Geweke diagnostics for convergence of variance components and heritabilities of BW, BD, and K were greater than 0.05. Since Geweke's statistic did not reveal any significant difference between the mean of the values sampled from the first and the last parts of the Markov chains, it indicates convergence and

Table 1. Estimates of variance components, heritability with their sampling variance and confidence intervals for various traits.

Variance	Trait - BW				Trait - BD				Trait – K			
	Estimate	SE	L - 95% CI	U - 95% CI	Estimate	SE	L - 95% CI	U - 95% CI	Estimate	SE	L - 95% CI	U - 95% CI
REML Animal Model												
Residual	325.20	31.90	262.70	387.70	0.0831	0.0048	0.0737	0.0925	9.97E-07	4.67E-08	9.06E-07	1.09E-06
Additive	255.20	51.68	153.90	356.50	0.0228	0.0056	0.0118	0.0338	1.12E-07	3.73E-08	3.85E-08	1.85E-07
Heritability	0.44	0.07	0.30	0.58	0.22	0.05	0.12	0.32	0.10	0.03	0.04	0.16
Animal Model Scheme – 1 (Non-Randomized NPBS with Independent u & e)												
Residual	240.00	28.31	188.10	299.20	0.0535	0.0067	0.0404	0.0666	8.71E-07	5.556E-08	7.621E-07	9.799E-07
Additive	230.00	24.00	182.60	277.10	0.033	0.005	0.0232	0.0428	2.09E-07	4.048E-08	1.300E-07	2.887E-07
Heritability	0.51	0.05	0.41	0.61	0.38	0.05	0.28	0.49	0.19	0.04	0.12	0.26
Animal Model Scheme – 2 (Non-Randomized NPBS with dependent u & e)												
Residual	248.41	36.14	177.58	319.24	0.0722	0.0085	0.0555	0.0889	9.56E-07	6.044E-08	8.380E-07	1.074E-06
Additive	358.52	31.65	296.45	420.55	0.0332	0.0057	0.022	0.0444	2.10E-07	4.300E-08	1.261E-07	2.947E-07
Heritability	0.41	0.05	0.31	0.51	0.32	0.05	0.21	0.41	0.18	0.04	0.10	0.25
Animal Model Scheme – 3 (Randomized NPBS with dependent u & e)												
Residual	302.46	37.04	229.16	375.76	0.0748	0.0088	0.0576	0.0920	9.74E-07	7.024E-08	8.363E-07	1.111E-06
Additive	311.80	33.80	245.60	378.05	0.0313	0.0059	0.0197	0.0429	2.01E-07	4.734E-08	1.086E-07	2.942E-07
Heritability	0.49	0.05	0.39	0.59	0.29	0.05	0.19	0.39	0.17	0.04	0.09	0.25
REML Family Model												
Residual	452.75	17.58	418.29	487.21	0.0946	0.0037	0.0874	0.1018	1.05E-06	4.10E-08	9.74E-07	1.13E-06
Family	128.24	25.92	77.44	179.04	0.0116	0.0028	0.0061	0.0172	5.72E-08	1.91E-08	1.98E-08	9.47E-08
Additive	256.48	51.84	154.87	358.08	0.0233	0.0057	0.0121	0.0344	1.14E-07	3.82E-08	3.96E-08	1.89E-07
Heritability	0.44	0.07	0.30	0.58	0.22	0.04	0.12	0.31	0.10	0.03	0.04	0.16
Family Model Scheme – 1 (Randomized NPBS with Independent u & e)												
Residual	103.60	16.40	71.50	135.70	0.0263	0.0073	0.0120	0.0406	1.07E-06	6.249E-08	9.523E-07	1.197E-06
Family	207.10	32.70	143.01	271.19	0.0016	0.0014	−0.0011	0.0043	5.39E-08	1.945E-08	1.576E-08	9.204E-08
Additive	498.97	26.40	447.23	550.71	0.0032	0.0027	−0.0021	0.0085	1.08E-07	3.891E-08	3.153E-08	1.840E-07
Heritability	0.34	0.05	0.24	0.44	0.11	0.07	−0.02	0.24	0.09	0.03	0.03	0.16
Family Model Scheme – 2 (Randomized NPBS with dependent u & e)												
Residual	133.60	17.90	98.52	168.68	0.0941	0.008	0.0784	0.1098	1.06E-06	5.432E-08	9.544E-07	1.167E-06
Family	267.20	35.70	197.23	337.17	0.0166	0.0028	0.0111	0.0221	1.07E-07	2.187E-08	6.407E-08	1.498E-07
Additive	524.40	23.04	479.24	569.56	0.0332	0.0056	0.0222	0.0442	2.14E-07	4.374E-08	1.281E-07	2.996E-07
Heritability	0.41	0.05	0.31	0.51	0.30	0.05	0.20	0.39	0.18	0.04	0.10	0.25
Family Model Scheme – 3 (Randomized NPBS with dependent u & e)												
Residual	145.40	16.70	112.67	178.13	0.0954	0.0083	0.0791	0.1117	1.07E-06	6.394E-08	9.495E-07	1.200E-06
Family	290.80	33.50	225.14	356.46	0.0149	0.0028	0.0094	0.0204	1.00E-07	2.313E-08	5.464E-08	1.453E-07
Additive	499.10	27.50	445.20	553.00	0.0298	0.0057	0.0186	0.0410	2.00E-07	4.627E-08	1.092E-07	2.906E-07
Heritability	0.45	0.05	0.35	0.55	0.27	0.05	0.18	0.37	0.17	0.04	0.09	0.25
Parametric Bootstrap												

Table 1. continued

Variance	Trait - BW				Trait - BD				Trait – K			
	Estimate	SE	L - 95% CI	U - 95% CI	Estimate	SE	L - 95% CI	U - 95% CI	Estimate	SE	L - 95% CI	U - 95% CI
Residual	325.20	31.90	259.70	384.50	0.0948	0.0055	0.0841	0.1055	9.95E-07	4.746E-08	9.018E-07	1.087E-06
Additive	255.60	51.50	162.60	363.60	0.0231	0.0060	0.0113	0.0349	1.13E-07	4.064E-08	3.297E-08	1.923E-07
Heritability	0.44	0.07	0.30	0.58	0.20	0.05	0.10	0.29	0.10	0.04	0.03	0.17
Bayesian MCMC												
Additive	279.40 (Mean) 273.00 (Median) 268.30 (Mode)	60.30	174.50	400.30	0.0257 (Mean) 0.0249 (Median) 0.0233 (Mode)	0.0068	0.0125	0.0389	0.00124 (Mean) 0.00120 (Median) 0.00114 (Mode)	0.00037	0.00057	0.00199
Residual	395.40 (Mean) 397.40 (Median) 404.80 (Mode)	37.40	321.20	467.00	0.1070 (Mean) 0.1069 (Median) 0.1068 (Mode)	0.0059	0.0953	0.1187	0.00992 (Mean) 0.00992 (Median) 0.00972 (Mode)	0.00045	0.00905	0.01082
Heritability	0.41 (Mean) 0.41 (Median) 0.43 (Mode)	0.07	0.28	0.55	0.19 (Mean) 0.19 (Median) 0.17 (Mode)	0.05	0.10	0.28	0.11 (Mean) 0.11 (Median) 0.10 (Mode)	0.03	0.05	0.17
Asymptotic Sampling												
Additive	257.60	51.20	164.60	365.10	0.0235	0.0058	0.0121	0.035	9.99E-07	4.628E-08	9.116E-07	1.092E-06
Residual	325.70	31.60	266.30	389.00	0.0832	0.0048	0.0739	0.0925	1.18E-07	4.018E-08	5.759E-08	2.136E-07
Heritability	0.44	0.07	0.30	0.58	0.22	0.05	0.12	0.31	0.10	0.03	0.05	0.18

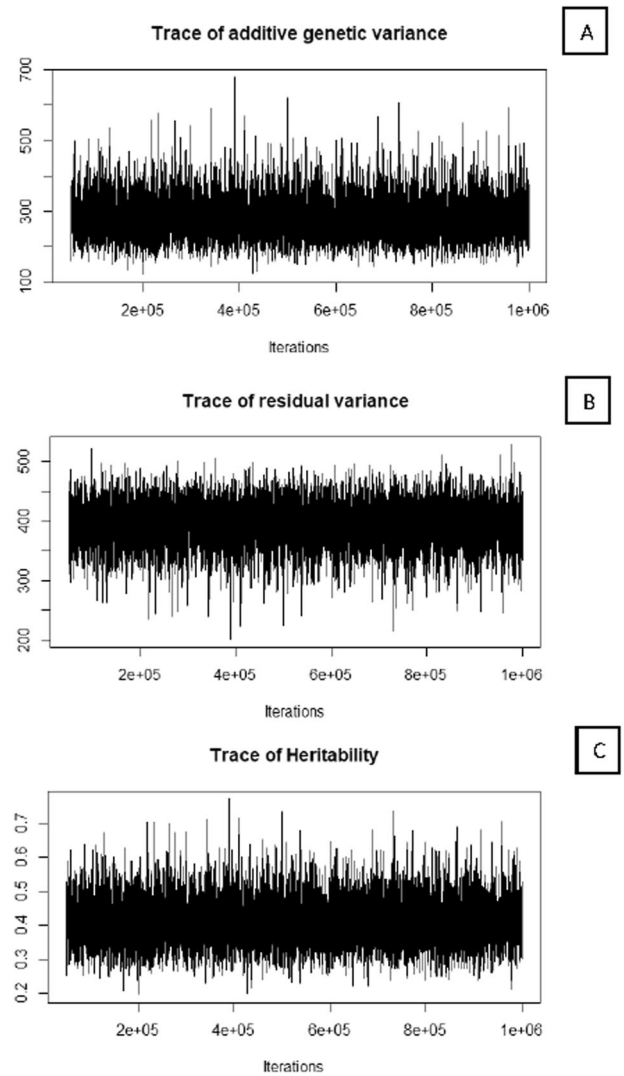

Fig. 1 Trace plots of the sampled values for additive variance, residual variance and heritability along the iterations from MCMC sampler for trait BW. (**A**) Additive genetic variance, (**B**) Residual variance, and (**C**) Heritability.

proper mixing of the chains. Also, the thinning interval of 100 reduced the autocorrelation of samples within the range of −0.001 to 0.10 for different traits (Table 2). Autocorrelations for additive genetic variances and residual variances for BW, BD, and K at five different points of iterations are given in Table 2, where Lag 100 stands for autocorrelation of estimates 100 iterations apart. The autocorrelation between additive genetic variance within the traits ranged from −0.001 to 0.10 and 0.001 to 0.05 for the residual variance.

High-density regions (HDRs). The posterior probability distribution of additive genetic variance, residual variance and heritabilities for trait BW are presented in Fig. 2 a, b. The additive genetic variance for BW ranged from 119.70 to 678.30 (See Fig. 2) with a 95% probability of the parameter values between 174.50 and 400.30, whereas the residual variance had 95% HDR ranging from 321.20 to 466.90 (Table 1). The posterior probability means of additive genetic variance and residual variance were 279.40, and 395.40 for BW, 0.0257 and 0.1070 for BD, 0.00124 and 0.00992 for K, and are presented in Table 1. The mean of the variance components and heritabilities for BW, BD and K are presented in Table 1. The heritabilities

Table 2. Autocorrelation values for different intervals for the MCMC chain.

	BW		BD		*K*	
Intervals	**Additive Variance**	**Residual Variance**	**Additive Variance**	**Residual Variance**	**Additive Variance**	**Residual Variance**
Lag 0	1	1	1	1	1	1
Lag 100	0.078	0.054	0.100	0.058	−0.011	0.001
Lag 500	0.007	0.007	−0.0001	0.001	−0.010	−0.023
Lag 1000	0.010	0.011	−0.0002	−0.002	−0.004	0.009
Lag 5000	−0.022	−0.005	−0.011	−0.004	−0.017	0.004

estimated by REML and the Bayesian method for BW, BD, and K were similar (Table 1).

Heritability estimates of BW obtained from the posterior distribution of variance components ranged from 0.20 to 0.77 with the posterior mean and median heritability of 0.41 and the posterior mode of 0.43, with a standard error of 0.07 for the mean heritability (Table 1). HDR 95% regions of heritability for BW ranged from 0.27 to 0.54, indicating a 95% probability for the true value of heritability to lie between 0.27 and 0.54. The mean heritability for BD and K were 0.19 and 0.11, respectively, which were similar to their median heritabilities.

Bootstrap estimates

Nonparametric bootstrap sampling (NPBS). The distribution of the variance components and heritabilities for BW obtained from nonparametric bootstrap is illustrated in Fig. 2 a, b. The distributions of variances and heritabilities of three traits are summarized in Table 1. A total of 10,000 bootstrap samples were obtained for each of the six scenarios, and the variance parameters were re-estimated by REML. Under the animal model, nonparametric bootstrap sampling was performed from the vector of animal effects (BLUPs) and residual effects sampling within fullsib (dam) families. The results are summarized in Table 1 with respect to the variance components and the heritability estimate. The 95% CI constructed for variance components and heritability by calculating the 2.5th and 97.5th percentile of bootstrap distribution is also given in Table 1.

The estimates of additive genetic variance (207.10–302.46), residual variance (Animal model: 230.00–358.52; Family model: 498.97–524.40) and heritabilities (0.34–0.51) for BW varied across the six NPBS schemes using animal model and family model, so as for BD and K (Table 1). For the NPBS models, resampling *u* and *e* independently without randomizing families (Scheme 1) gave rise to the lowest estimated residual variances for BW, BD, and K under the animal model; under the family model, scheme 2 had the lowest residual for K. Recognizing the dependence (correlation) between the estimates of and the randomization of families (Scheme 3) yielded the highest estimated residual variance for all traits under both animal model and family model. Assuming independent *u* and *e* in the animal model (Scheme 1) has inflated heritability (the quantum of inflation increased with decreasing REML heritability), and under dependent *u* and *e* assumption in the animal model, the non-randomization (Scheme 2) and randomization (Scheme 3) yielded heritability estimates with differences of the order 0.08 for BW (Scheme 2: 0.41; Scheme 3: 0.49), 0.03 for BD (Scheme 2: 0.32; Scheme 3: 0.29) and 0.01 for K (Scheme 2: 0.18; Scheme 3: 0.17). The family variance was the lowest under the independent *u* and *e* model even when the families were randomized (Scheme 1) for all three traits. Under the dependent *u* and *e* assumption in the family model, the family randomization (Scheme 3) gave higher family variance for BW. For BD and K, the non-randomization (Scheme 2) and the randomization (Scheme 3) yielded similar family variance. For both the animal and family model, the highest residual variance was observed under the dependent *u* and *e* model with family randomization for all three traits. Assuming independence of *u* and *e*, family model scheme 1 gave the lowest heritability for all three traits (Table 1). Additionally, the results from various NPBS schemes, applied on different data sets with varying family size and the total sample size randomly sampled from the original data, is attached as a supplementary file (Annexure-2).

Parametric bootstrap. The results obtained from the parametric bootstrap are presented in Table 1. The distribution of variance components and heritabilities for BW are illustrated in Fig. 2. The estimates of variance components were calculated as the mean of the parameter estimates from bootstrap samples and the confidence interval as the values lying between 2.5th and 97.5th percentile. The additive genetic variance and residual variance for different traits estimated from the parametric bootstrap, along with its standard errors, are provided in Table 1. The variance and heritability estimates for three traits obtained from the parametric bootstrap were similar to that of the REML estimates (Table 1).

Asymptotic sampling. The estimate of variance components and heritabilities for different traits were obtained as the mean of all the sampled values and is presented in Table 1. The distribution of the sampled values for additive genetic effects and residual effects and heritabilities for BW is given in Fig. 2 a, b. The values of variances and heritabilities estimated for the traits considered were similar to REML estimates.

Comparison of estimates. Variance components, heritabilities and their uncertainties were estimated from five different methods, of which the heritabilities ranged from 0.34 to 0.51 for BW, 0.11 to 0.38 for BD and 0.09 to 0.19 for K across the methods (Table 1). Estimates of heritabilities and standard errors obtained from REML, parametric bootstrap and asymptotic sampling were similar for different traits (BW: 0.44, BD: 0.22 and K: 0.10). MCMC sampling also gave similar heritability estimates as REML for different traits. Under the NPBS models, the average heritability of BW for Scheme 3 was 0.49 ± 0.07, less than 1 SD above the REML estimate. The other schemes disagreed more with the REML estimate, but the REML estimate was still well within the 95% coverage range. The 95% coverage probabilities of variance components and heritabilities from different methods for BW are given in Table 3. The coverage probabilities obtained for different methods ranged from 0.94 to 0.98 (Table 3). For the traits BD and K, the REML heritability was within the 95 % CL of the NPBS scheme 3 using both animal model and family model, and the largest bias was observed in scheme 1.

For subjective comparison, overlaid kernel density graphs for additive variance, residual variance and heritabilities of BW obtained from different sampling-based methods are illustrated in Fig. 3 a–c. These figures show that the distribution of additive genetic variance, residual variance and heritabilities are approximately normal. The location of the peak of distribution for the additive genetic variance does not vary significantly (well within 1 SD of REML estimate of additive variance) among the methods, with the smallest sampling variance for the additive genetic

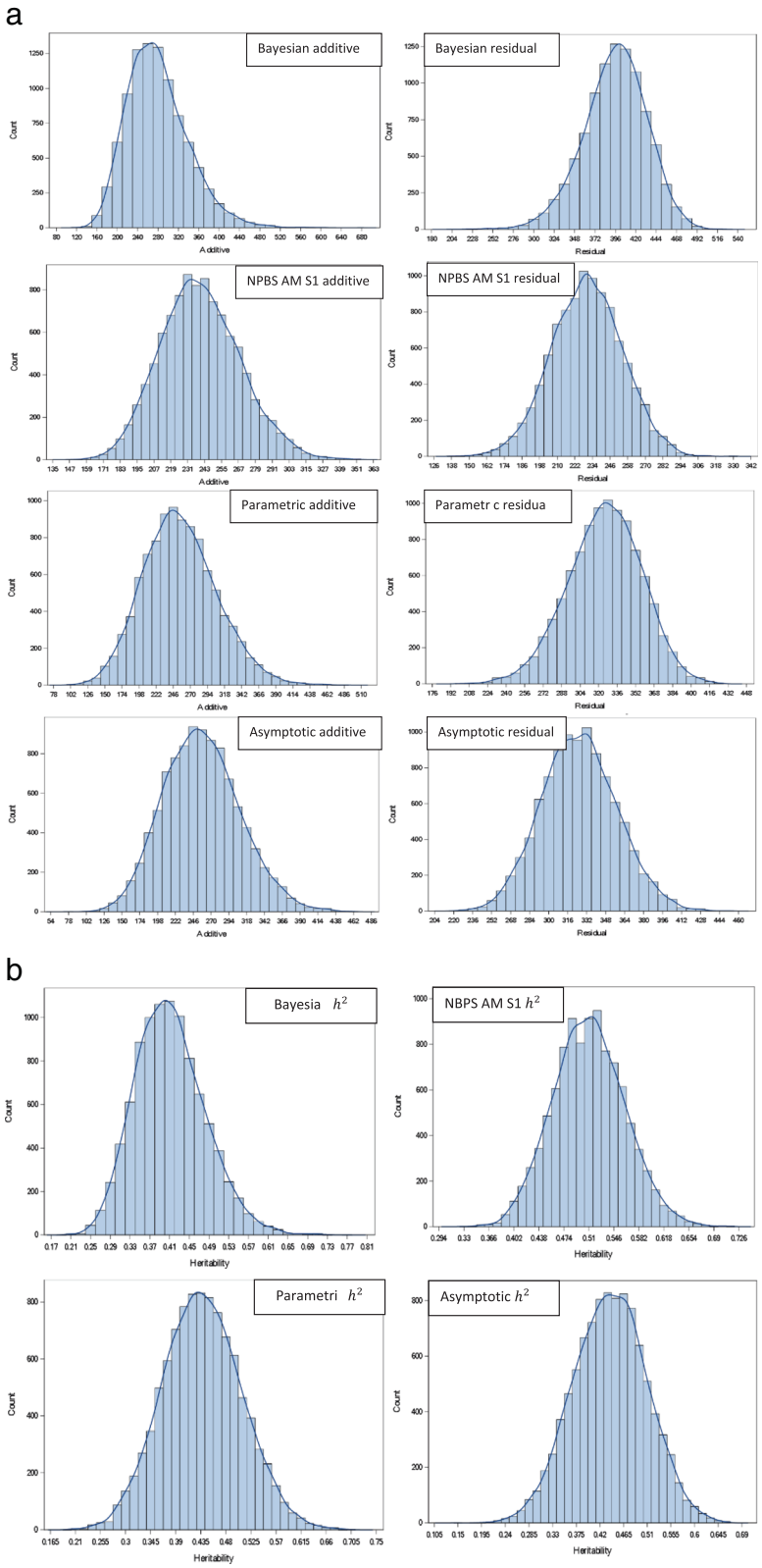

Fig. 2 Distribution of variance components and heritabilities estimated for BW from four methods. a A (Column 1: Additive variance; Column 2: Residual variance; Row 1: Bayesian MCMC, Row 2: NPBS AM scheme 1, Row 3: Parametric bootstrapping and Row 4: Asymptotic sampling. **b** Heritabilities Row1-Column1: Bayesian MCMC, Row1-Column2: NPBS AM scheme 1, Row2-Column1: Parametric bootstrapping and Row2-Column2: Asymptotic sampling.

Table 3. Coverage probabilities for NPB: Nonparametric bootstrap, PB: Parametric bootstrap, BYS: Bayesian estimation, ASY: Asymptotic sampling.

	NPBS	PB	BYS	ASY
Additive	0.98	0.94	0.96	0.95
Residual	0.97	0.94	0.97	0.94
Heritability	0.97	0.94	0.96	0.95

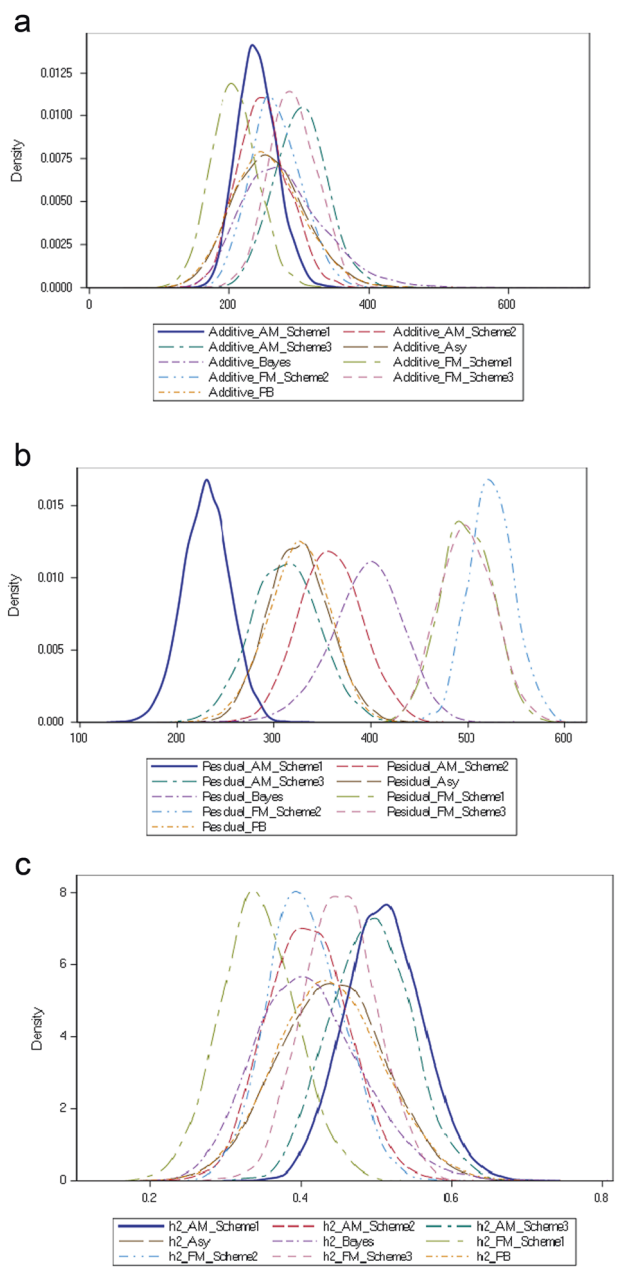

Fig. 3 Overlaid kernel density graphs for BW obtained from different methods. a Additive genetic variances, **b** Residual variances, and **c** Heritabilities.

variance obtained from NPBS models (with a sharp peak and narrow confidence interval) giving higher confidence in the parameter estimated. Among different methods, the sampling variance for additive variance was high for the MCMC method with a wider confidence interval. The distribution for residual variance varied largely among methods (Fig. 3b), where both the lowest and the highest residual variance were obtained from NPBS models. Both the parametric bootstrap and asymptotic sampling gave rise to similar distribution for a residual variance. The sampling variance obtained from REML, parametric bootstrap, MCMC method and asymptotic sampling for a residual variance were similar. The parametric bootstrap, MCMC method and asymptotic sampling gave identical distribution for heritability estimates. The distribution for the NPBS model with independent *u* and *e* assumption largely varied from the empirical distribution of other methods (Fig. 3c).

DISCUSSION

In REML estimation, the maximization of likelihood can be achieved differently depending upon the order of derivatives available (Thompson and Mäntysaari 1999). The precise class of algorithms are derivative-free (DF) methods, of which simplex/polytope (Nelder and Mead 1965) and Powell's algorithms (Powell 1964) are more common. The most popular algorithm involving the first derivatives is Expectation-Maximization (EM REML) algorithm, while the Average-Information algorithm (AI REML) is the most popular involving second derivatives (Misztal 2008).

The DF approaches are fast for simple models, whereas they are expensive and unreliable for complicated models and have become obsolete (Gilmour et al. 1995). The EM algorithm was considered the most reliable but is slow in convergence and does not directly generate the standard error of the estimate. AI REML is popular because it is much easier to compute than other second derivative methods like Newton-Raphson and Fisher Scoring, it is not much more complicated than EM per iteration and it requires many fewer iterations (Gilmour et al. 1995). One advantage of using the Newton-Raphson or Average Information algorithm is that the matrix of the second derivatives of the log-likelihood evaluated at the optima, known as the Hessian matrix (H), is available upon completion. Serfling (1980) explained with the help of the asymptotic theory of maximum likelihood that matrix $2H^{-1}$ is an asymptotic variance-covariance matrix of the estimated parameters of G and R. In the breeding experiments, heritability is the primary genetic parameter used to assess the potential for genetic improvement. It is estimated as a non-linear function (ratio) of variance components (Meyer and Houle 2013; Stefan 2017). Under the REML method, the estimated heritability's reliability/standard error (SE) is obtained from the first-order Taylor series approximation of the function of heritability. In all these methods, the estimation of reliabilities is based on the large sample (asymptotic) theory, where the test and confidence intervals are based on the asymptotic normality (Serfling 1980). These estimates of standard errors may be unreliable, especially under small sample sizes, since these methods are based on approximations (Littell et al. 2006). In the present study, we compared the variance components and their reliabilities and the precision of heritability estimates obtained from the REML method with the corresponding estimates obtained using various sampling methods.

The sampling-based methods are costly on time compared to the REML method. However, the time required for sampling-based methods depends upon the sample size, the number of iterations envisaged, and the software and the computational capacity. The recorded CPU time ranged from 45 to 90 min in the present study, depending upon the method (AMD A10-8700P Radeon R6, 10 Compute Cores 4 C+6 G 1.80 GHz with 8 GB installed RAM).

Bootstrap vs REML estimates

Bootstrap methods are an alternative approach for estimating the reliability/SE of parameters and constructing the confidence interval without assuming the symmetric distribution. Model-based bootstrapping was first developed under simple linear

models, where bootstrap samples were based on residuals (Efron 1979; Efron and Tibshirani 1994). In the simple linear models, the residuals are resampled either from an estimated empirical distribution of the residuals (parametric bootstrap) or are resampled from the initial residuals without any distributional assumptions (nonparametric bootstrap). The resampled residuals are then added onto an estimate of the mean function obtained from the data (Morris 2002) to create a new set of observations to analyze. In the present study, the same idea was extended to the animal model by sampling with replacement from predictors of the random effects and residuals for nonparametric bootstrap and from an estimated multivariate normal distribution for parametric bootstrap. In the context of an animal model, the natural choice of predictors for random effects is BLUPs (BLUP bootstrap), as they are readily available in software packages and have optimality properties in predicting an individual's random effect. However, the variance structure is defined by the sum of the random effect and the residual. A random effect is a shrunken form of a fixed effect. The balance of the fixed effect remains in the residual, so resampling needs to be done on the sum to reproduce the appropriate variance structure.

Different NPBS models were used in this study. The NPBS with an animal model assuming independence of u and e (Scheme 1) gave rise to a high heritability score for all the three traits (For BW REML: 0.44, Scheme 1: 0.51; BD REML: 0.22, Scheme 1: 0.38; and K REML: 0.10, Scheme 1: 0.19) wherein all three cases the residual variance was underestimated relative to REML. This suggests this bootstrapping process is not recreating the variance structure assumed for the model, and it is because the estimated random effect is correlated with the residual because of the shrinkage in its estimation. Combining an animal effect ($(f_i + e_{ij})s$) with the residual from another animal in the same family ($(f_i + e_{ik})(1\text{-}s)$) produces ($f_i + se_{ij} + (1\text{-}s)e_{ik}$) in which the latter two terms tend to cancel producing a smaller residual than required. The math is different under the family model bootstrapping, where under scheme 1 the family variance is low, and the residual is high: the family effect ($s_i\ f_i$) is combined with a residual ($1\text{-}s_k$) $f_k + e_{kj}$ typically from another family.

Under Schemes 2 and 3, adding the random and residual components at the individual level always yields a sum that is not disturbed by the shrinkage or whether it is fitted as an animal or family model. However, the average sampled genetic variance was inflated relative to the initial REML estimates in both cases for all traits, with a commensurate reduction in the residual variance. The REML solution has the same standard deviation and is well within 1 standard deviation (SD) of the mean under Scheme 3 for BW, within 2 SD for BD and K. However, the failure to randomize family effects under scheme 2 has resulted in a family variance 2 SD above the REML estimate. This indicates that Scheme 2 is not properly representing the variation in the data.

A simulation study by Morris (2002) demonstrated that the optimal properties of the BLUP do not transfer over to bootstrapping when the sample size is small, as a result of which the BLUP based bootstrap consistently underestimate the variability in the data. Further, he reported that the coverage probability of 90 % intervals for the variance components from BLUP bootstrap showed severe undercoverage problems. However, there were no undercoverage issues in the present study with the 95% confidence intervals for variance components and heritability, which add confidence to the optimality of the current sample size (Table 3). An increase in the coverage rate with an increase in sample size was reported by Thai et al. (2013). In our study, both nonparametric and parametric bootstrap gave similar coverage probabilities with slight differences in the values attributed to the total bootstrap sample size (10,000). Increasing the total bootstrap sample size might give exact coverage probability for nonparametric and parametric bootstrap. Searle et al. (2009) explained that the realizations of BLUP are shrinkage estimates and the overall effect of shrinkage estimation is the reduction in the variance surrounding realized BLUPs (Morris 2002). This reduction in variance is translated to every new data set generated by resampling from BLUP predictions. In a mixed model analysis with pedigree, the BLUP shrinkage is a function of the number of observations per family, the total number of observations, the observed vector of trait values and estimated variance components (Searle et al. 2009). So, one of the possible ways to reduce the effect of shrinkage would be to randomize the families while bootstrapping so that the overall shrinkage effect might cancel out.

The heritability estimated from the parametric bootstrap was similar to the REML estimate. In parametric bootstrap, the random variables were sampled from a multivariate normal distribution, resulting in an estimate identical to the REML method since the latter is based on the strong assumption of multivariate normality. A global nonparametric bootstrap was also performed (results not shown), in which the random animal effects and residuals were sampled from the respective vectors of BLUPs and residuals, wherein each vectors were respectively assumed to be one single sampling unit in contrast to assuming each fullsib family as the primary sampling unit. The global nonparametric bootstrap did not yield any meaningful result, where the estimated variance was near zero. The reason for this is, while performing bootstrapping, one of the concerns is that the resampling should appropriately mimic the actual data generating process that produced the data set (Flachaire 2005). It is evident that the classical bootstrap methods developed for simple linear models should be modified to take into account the characteristics of mixed-effects models (Das and Krishen 1999). In the setting of an animal model, the within-family correlation structure needs to be taken into account; thereby, one parent and its entire offspring will be the appropriate sampling unit for generating meaningful bootstrap samples, rather than a whole vector of animal effects (predicted breeding values).

Bayesian vs REML estimates

Bayesian methods are often touted for their ability to incorporate prior information when available, but the key utility of these methods is their ability to provide a complete description of the uncertainty of an estimate (Walsh and Lynch 2018). An MCMC chain of 10,00,000 iterations were run with a burn-in of 50,000 iterations (for better convergence and mixing of chain) and a thinning interval of 100 (to reduce autocorrelation), which yielded a total of 9500 sampled values of variance components and heritabilities. Convergence and autocorrelation are two critical issues to monitor when using the MCMC method. There is a possibility of a strong dependence of the values obtained in the first few iterations on the starting values. The chain is said to converge only after the dependence on the starting parameter has diminished. The visual examination of the trace suggested a well-mixing chain (Fig. 1). As a rule of thumb, if there is no trend in the trace, then the chain has achieved convergence (Hadfield et al. 2019). An autocorrelation of less than 0.10 in magnitude is considered reasonable (de Villemereuil 2012). Further, if the samples were drawn from the chain's stationary distribution, the two means are equal, and Geweke's statistic has an asymptotically standard normal distribution (Plummer et al. 2018).

One of the consequences of the large sample theory is the asymptotic normality of the posterior distribution, i.e., to say as $n \rightarrow \infty$ (as more and more data arrive from the underlying process), the posterior distribution of θ approaches normality (Gelman et al. 2013). Examination of the distribution of the variance components (additive and residual; presented for BW in Fig. 2a) for different traits obtained from the MCMC sampler shows that the posterior distribution is approximately normal. The additive variance obtained from the REML analysis was in close agreement with the location parameters of the posterior

distribution, whereas the REML residual variance was notably less in comparison with the Bayesian posterior estimate for BW and BD; however, for trait K in its original scale, the MCMC did not yield results similar to REML solutions which inflated the heritability. The probable reason could be that a weakly informative inverse gamma prior was used to obtain the posterior distribution. In the original scale, the likelihood estimates of variance components of K is of the order 10^{-7} and the parameterization of inverse gamma prior used in this study has its density function approaching infinity as the variance approaches zero. It means the information of the prior is maximum as the variance approached zero, or in other words, the only information available in the prior is that variance cannot be negative. The high information of prior near to zero will always overpower the information in the likelihood estimates as it approaches zero, which was the case with K (likelihood estimate of the order of 10^{-7}). To remove this inconvenience, the observations of K were multiplied with an appropriate constant (100 in the present study), which yielded variance components of the order 10^{-7}, the ratios of which estimated the heritability similar to REML. The scale corrected *K* values were used only in the Bayesian method. Gelman (2006) noted the popularity of the inverse gamma family of priors is due to its clean mathematical properties, wherein he discourages the use of inverse gamma prior due to the sensitivity of its parameterization towards the posterior inferences. However, in the present study, the MCMCglmm package was used, which does not offer the flexibility to use a different family of prior other than inverse gamma. Waldmann and Ericsson (2006) reported a high residual variance in the posterior distribution when comparing the REML and Gibbs sampling estimates of genetic parameters, which was the case with BW and BD in this study. However, according to Sorensen et al. (2002), the variance components from REML analysis should be identical to the Bayesian posterior distribution mode if the mixed models' parameters are assigned non-informative uniform distributions. Another study by Guan et al. (2017) in turbot (a species of flatfish) showed that the Bayesian estimate of additive variance from the posterior mean was high compared to the REML estimate, whereas the Bayesian residual variance was low in comparison with the corresponding REML estimate. In the present study, a higher additive and residual variance were obtained from the posterior mean, median, and mode compared to the REML estimates for BW and BD, whereas for K, the MCMC variances agreed with REML solutions (Additive: 0.0011 and Residual: 0.0099). In a simulation study by Van Tassell and Van Vleck (1996), the variance components obtained from the posterior mean and REML estimates were similar, and the variance components obtained from the posterior mode was always low. Despite the difference in variances, the posterior distribution of heritability resulted in a similar value as that of the REML estimate, giving more confidence in our estimate. Several other studies also reported where both REML and Bayesian methods provide a similar estimate of heritability if the sample size is large and heritability is high (Waldmann and Ericsson 2006; Alijani et al. 2012; de Villemereuil et al. 2013; Guan et al. 2017). It should be noted that the standard error estimated from the posterior distribution was similar to the one obtained from the delta method. The approximate 95% CI of REML heritability agreed well with the 95% HDR of heritability of posterior distribution. An objective comparison between the estimates of the REML and Bayesian posterior distribution is difficult. However, heritabilities obtained from both the methods were in close agreement, which indicates that the likelihood function has an overpowering influence on the prior distribution. The influence of the prior distribution on the posterior diminishes under a large sample size (Walsh and Lynch 2018); nevertheless, only a prior sensitivity study can ascertain this claim, as noted by Blasco and Blasco (2017).

Asymptotic sampling vs REML estimates

Meyer and Houle (2013) described a simple alternative to estimate the sampling distributions of the functions of variance components by repeated sampling of parameter estimates from their asymptotic, multivariate normal (MVN) distribution and calculating the functions of interest for each sample inspecting their distribution across replicates. They further concluded that the sampling of REML estimates from asymptotic MVN distribution, specified by the inverse of the information matrix, offers a straightforward and computationally undemanding way to derive sampling distributions and confidence intervals for estimates of covariance components and their functions. Our study sampled 10,000 estimates of both additive genetic and residual variances for all three traits and estimated heritability for the combination of every sampled value. The mean of the variance components and heritabilities obtained from the asymptotic sampling were similar to the REML estimates for different traits. The standard error for variance components and heritabilities obtained from both methods were similar. In the present study, both the REML and asymptotic sampling performed equally well to estimate variance components' uncertainties and functions. An admonition in sampling from asymptotic distribution is that, to yield a valid estimate of covariance, sampling distributions and confidence intervals, large sample properties should hold well, i.e., the inverse of the information matrix has to provide an adequate description of sampling covariance among the parameters estimated. The inverse of the information matrix could be obtained only with Newton–Raphson type algorithms or its variant, especially the average information algorithm (AIREML) (Gilmour et al. 1995), which utilizes second derivatives of log-likelihood.

The uncertainties surrounding the REML heritability could be inaccurate when the assumed asymptotic behavior is violated. Schweiger et al. (2016) describe different cases under which the asymptotic normality does not hold well, of which the most common cause is the sample size. They further demonstrated that the asymptotic CIs tend to be biased when the heritability estimated is relatively low or high, under which the CIs can spread beyond the natural boundaries of their parameters (e.g., negative heritabilities). In a breeding program, the number of families and the number of offspring per family that could be generated at a time is constrained by many factors. As we have noted previously, under large sample assumptions, it is not clear what constitutes a large sample (Walsh and Lynch 2018), making it hard to realize the goodness of asymptotic approximation. As an alternative, sampling-based approaches can be used to generate a full distribution of the estimates to construct CIs.

Of the alternate methods used in the present study, the parametric bootstrap and asymptotic sampling gave similar results for variance components and heritability as REML estimates for all the traits under study. Even though the methods were different, the underlying assumptions are the same and give rise to similar results. Nonparametric bootstrap is well-known for estimating the uncertainties in the data by generating a distribution of estimates. In our study, we have used the BLUP predictions for resampling to create new datasets. However, BLUP is a shrinkage estimator, and as a result, there can be a bias in the estimate of heritability; however, the bias can be reduced by sampling under appropriate assumptions. The NPBS scheme 3 has minimum assumptions, and also, the confidence interval constructed by this method encompassed the heritability estimated by the REML method. The Bayesian posterior mode for BW, BD, and K yielded heritability estimates similar to REML heritability; however, there were slight differences in the variance estimated from both methods.

We conclude that this study with 1413 observations representing 78 full-sib families has provided sufficient information to estimate the heritability of the traits analyzed with acceptable confidence using REML. Based on our study, we recommend to use the NPBS scheme 3 or Bayesian methods to estimate the heritability and its standard error when the information content of the data is in doubt.

CONCLUSION

Variance components and heritabilities are point estimates surrounded by statistical uncertainty. Confidence intervals derived from the standard error of the estimate describe the uncertainty of the estimate. The REML method assumes an underlying normal distribution leading to asymptotic normality of the estimators, but this may not hold for small sample sizes. When it is challenging to ascertain adequate sample sizes, it is possible to use sampling-based methods to generate a full distribution of estimates. In the present study, the REML estimates of variance components, heritability and the associated uncertainties for three traits were compared with different sampling-based approaches to understand if the data had sufficient information for the asymptotic assumptions to hold.

Even though the parametric bootstrap and asymptotic sampling yielded precisely the same variance components for all three traits as those estimated from the REML method, these two methods also assume multivariate normality under a large sample theory like REML, hence not the methods of the first choice. Moreover, the second moment of the assumed multivariate normality from which the parametric bootstrap and asymptotic sampling were performed was the REML estimate of variance components. The NPBS estimates vary based on the assumptions, where the stronger the assumption, the tighter was the distribution. Based on the NPBS results, obtaining the bootstrap sample by independently resampling the genetic and residual components is not valid because of the effect of shrinkage on the genetic effects meaning their estimates are correlated with the residual effects. However, it is necessary to randomize the family effects to sample the parameter space adequately. The REML solution for BW was within 1 SD, and BD and K within 2 SD of the bootstrap mean based on the randomization of family effects, indicating strong confidence in REML solutions. The heritability estimated from the mode of Bayesian posterior density did not deviate from the REML heritability for all traits. The randomized NPBS with linked genetic and residual effects is the sampling method with the least assumptions. To ascertain the adequacy of the sample size to estimate reliable genetic variation, we recommend NPBS scheme 3 as it provides high overlap with REML calculations and requires the least assumptions. The present study shows that the heritability estimated from different methods and schemes are similar to the REML estimates and the confidence intervals are of similar magnitude and largely overlapping. Hence, it is concluded that the present data set meets the assumptions made for likelihood analysis, and the heritability is well estimated by REML.

DATA AVAILABILITY

The codes used for the analysis of the data employing R, Wombat and Echidna are provided as supplementary material. For access to the raw data, please contact the corresponding author.

REFERENCES

Alijani S, Jasouri M, Pirany N, Kia HD (2012) Estimation of variance components for some production traits of Iranian Holstein dairy cattle using Bayesian and AI-REML methods. Pak Vet J 32(4):562–566

Blasco A, Blasco PDA (2017) Bayesian data analysis for animal scientists. Springer, New York, NY, USA, (Vol. 265)

Das S, Krishen A (1999) Some bootstrap methods in nonlinear mixed-effect models. J Stat Plan Inference 75(2):237–245

de Villemereuil P (2012) Estimation of a biological trait heritability using the animal model. *How to use the MCMCglmm R package*, 1-36

de Villemereuil P, Gimenez O, Doligez B (2013) Comparing parent–offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits. Methods Ecol Evolution 4(3):260–275

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26

Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. CRC press, London, UK

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longmans Green, Harlow, Essex, UK

Flachaire E (2005) Bootstrapping heteroskedastic regression models: wild bootstrap vs. pairs bootstrap. Computational Stat Data Anal 49(2):361–376

Gelman A (2006) Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal 1(3):515–534

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian data analysis. CRC press, Broken Sound Parkway NW, Suite 300 Boca Raton, FL

Geweke J (1992) Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments Bayesian Stat 4:641–649

Gilmour AR (2018) *Echidna Mixed Models Software*. Paper presented at the *proceedings of the world congress on genetics applied to livestock production*, Auckland, New Zealand. http://www.wcgalp.org/proceedings/2018

Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R (2015) ASReml User Guide Release 4.1 Structural Specification, *VSN International Ltd*, Hemel Hempstead, HP1 1ES, UK www.vsni.co.uk

Gilmour AR, Thompson R, Cullis BR (1995) Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. Biometrics, 1440-1450

Guan J, Wang W, Hu Y, Wang M, Tian T, Kong J (2017) Estimation of genetic parameters for growth trait of turbot using Bayesian and REML approaches. Acta Oceanologica Sin 36(6):47–51

Hadfield J, Hadfield MJ, SystemRequirements C (2019) Package 'MCMCglmm'. See https://cran.rproject.org.

Henderson CR (1973) Sire evaluation and genetic trends. J Anim Sci 1973 (Symposium):10–41

Isik F, Holland J, Maltecca C (2017) Genetic data analysis for plant and animal breeding. Springer International Publishing, Cham, Switzerland, (Vol. 400)

Jousy N, Jahageerdar S, Prasad JK, Babu PG, Krishna G (2018) Body weight at harvest and its heritability estimate in Clarias magur (Hamilton, 1822) reared under mono and polyculture systems. Indian J Fish 65(2):82–88

Kruuk LE (2004) Estimating genetic parameters in natural populations using the 'animal model'. Philos Transactionsof R Soc Lond 359(1446):873–890. Series B: Biological Sciences

Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Oliver S (2006) SAS for mixed models. SAS institute, Cary, NC, SAS publishing. (Vol. 2)

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA, Vol. 1

Madsen P, Jensen J, Labouriau R, Christensen OF, Sahana G (2014, August) DMU-a package for analyzing multivariate mixed models in quantitative genetics and genomics. In *Proceedings of the 10th world congress of genetics applied tolivestock production* (pp. 18–22)

Meyer K (2007) WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J Zhejiang Univ Sci B 8(11):815–821

Meyer K, Houle D (2013, October) Sampling based approximation of confidence intervals for functions of genetic covariance matrices. In *Proc. Assoc. Advmt. Anim. Breed. Genet* (Vol. 20, pp. 523–526)

Misztal I (2008) Reliable computing in estimation of variance components. J Anim Breed Genet 125(6):363–370

Morris JS (2002) The BLUPs are not "best" when it comes to bootstrapping. Stat Probab Lett 56(4):425–430

Nelder JA, Mead R (1965) A simplex method for function minimization. Computer J 7 (4):308–313

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58(3):545–554

Plummer M, Best N, Cowles K, Vines K (2018) Package 'coda'

Powell MJ (1964) An efficient method for finding the minimum of a function of several variables without calculating derivatives. Computer J 7(2):155–162

Psutka JV, Psutka J (2019) Sample size for maximum-likelihood estimates of gaussian model depending on dimensionality of pattern space. Pattern Recognit 91:25–33

Rameez R, Jahageerdar S, Chanu TI, Jayaraman J, Bangera R (2020) Genetic variation among full-sib families and the effect of non-genetic factors on growth traits at harvest in Clarias magur (Hamilton, 1822). Aquac Rep. 18:100411

Robinson GK (1991) That BLUP is a good thing: the estimation of random effects. Stat Sci 6(1):15–32

Schweiger R, Kaufman S, Laaksonen R, Kleber ME, März W, Eskin E, Halperin E (2016) Fast and accurate construction of confidence intervals for heritability. Am J Hum Genet 98(6):1181–1192

Searle SR, Casella G, McCulloch CE (2009) Variance components. John Wiley & Sons, Ney Jersey, USA

Serfling RJ (1980) Approximation theorems of mathematical statistics. John Wiley and Sons, New York, NY

Sorensen D, Gianola D, Gianola D (2002) Likelihood, Bayesian and MCMC methods in quantitative genetics. Springer, New York, USA

Stefan ME (2017). Standard errors of heritability estimates. Online (http://www.iysik.com/dmu/heritabilities)

Thai HT, Mentré F, Holford NH, Veyrat-Follet C, Comets E (2013) A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models. Pharm Stat 12(3):129–140

Thompson R, Mantysaari EA (1999) Prospects for statistical methods in dairy cattle breeding. *Proceedings of the Genetic Improvement of Functional Traits in cattle longevity, Interbull. Jouy-en-Jossas*, France, 1999, 20, 71–78

Van Tassell CP, Van Vleck LD (1996) Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co) variance component inference. J Anim Sci 74(11):2586–2597

Waldmann P, Ericsson T (2006) Comparison of REML and Gibbs sampling estimates of multi-trait genetic parameters in Scots pine. Theor Appl Genet 112(8):1441–1451

Walsh B, Lynch M (2018) Evolution and selection of quantitative traits. Oxford University Press, New York, USA

ACKNOWLEDGEMENTS

This study formed part of the Ph. D. program of the first author and he acknowledges the financial support received in the form of a Senior Research Fellowship from ICAR, New Delhi, India. The authors like to thank the Director, ICAR-CIFE, for providing the facility to carry this study. The authors extend their gratitude to SIC, technical and supporting staff of Freshwater Fish Farm, CIFE Center, Balabhadrapuram/Kakinada, Andhra Pradesh, India.

AUTHOR CONTRIBUTIONS

RR: Field experiment, data standardization, analysis, preparation of draft, programing algorithms. SJ: Conceptualization, experimental procedures, methodology, resources and fund acquisition, validation of formal analysis, review and editing, fine tuning. JJ: Methodology and validation, algorithm design and programming, review and editing. TIC: Aquaculture and reproduction management, sampling and data recording, and editing. RB: methodological inputs and validation, review and editing, AG: Theoretical considerations of different algorithms, validation of algorithms, tuning of algorithms and interpretation of results, review and editing.

COMPETING INTERESTS

The authors declares no competing interests.

ANIMAL ETHICS

The present experiment was part of a research project approved by the Institute Research Council (IRC) of ICAR- Central Institute of Fisheries Education (Deemed University), Mumbai, India (Project code: CIFE-2012/7). The experimental procedures followed were in compliance with the guidelines of IRC.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-022-00509-1.

Correspondence and requests for materials should be addressed to Shrinivas Jahageerdar.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.