
Genetics and population analysis

Estimation of dynamic SNP-heritability with Bayesian Gaussian process models

Arttu Arjas¹, Andreas Hauptmann^{1,2} and Mikko J. Sillanpää^{1,3,*}

¹Research Unit of Mathematical Sciences, University of Oulu, Oulu, FI-90014, Finland

²Department of Computer Science, University College London, London, UK

³Infotech Oulu, University of Oulu, Oulu, FI-90014, Finland

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Improved DNA technology has made it practical to estimate single nucleotide polymorphism (SNP)-heritability among distantly related individuals with unknown relationships. For growth and development related traits, it is meaningful to base SNP-heritability estimation on longitudinal data due to the time-dependency of the process. However, only few statistical methods have been developed so far for estimating dynamic SNP-heritability and quantifying its full uncertainty.

Results: We introduce a completely tuning-free Bayesian Gaussian process (GP) based approach for estimating dynamic variance components and heritability as their function. For parameter estimation, we use a modern Markov Chain Monte Carlo (MCMC) method which allows full uncertainty quantification. Several data sets are analysed and our results clearly illustrate that the 95 % credible intervals of the proposed joint estimation method (which "borrows strength" from adjacent time points) are significantly narrower than of a two-stage baseline method that first estimates the variance components at each time point independently and then performs smoothing. We compare the method with a random regression model using MTG2 and BLUPF90 softwares and quantitative measures indicate superior performance of our method. Results are presented for simulated and real data with up to 1000 time points. Finally, we demonstrate scalability of the proposed method for simulated data with tens of thousands of individuals.

Availability: The C++ implementation dynBGP and simulated data are available in GitHub (<https://github.com/aarjas/dynBGP>). The programs can be run in R. Real datasets are available in QTL archive (<https://phenome.jax.org/centers/QTLA>).

Contact: mikko.sillanpaa@oulu.fi

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Heritability, the proportion of phenotypic variation attributable to genetic factors, is a fundamental parameter in population and quantitative genetics (Visscher *et al.*, 2008). The aim in narrow-sense heritability estimation is to separate the variance of the trait into additive genetic and environmental variance components such that their sum equals the total variance of the trait. Generally, heritability is a population-specific parameter which can be estimated either using (i) linear mixed model (LMM) techniques (Henderson, 1984) or with (ii) multi-locus association (MLA) approaches

(Sillanpää, 2011). In LMMs, information of the (additive) genetic relationships between the individuals in the population must be available, which can be determined either from known pedigree or from genomic data (single nucleotide polymorphism, SNP), while in MLA models, heritability can be estimated from genomic data. In LMMs, the trait variation is assumed to be controlled by a high number of small effect genes (polygenic), whereas MLA models assume that there are only few major genes behind the trait variation. We note, that in case heritability is estimated from genomic data, it is called SNP-heritability, and we will from now on refer by heritability specifically to narrow-sense SNP-heritability.

It is well known that heritability of a trait may depend on measurement time, age, environmental conditions (like temperature), or

1

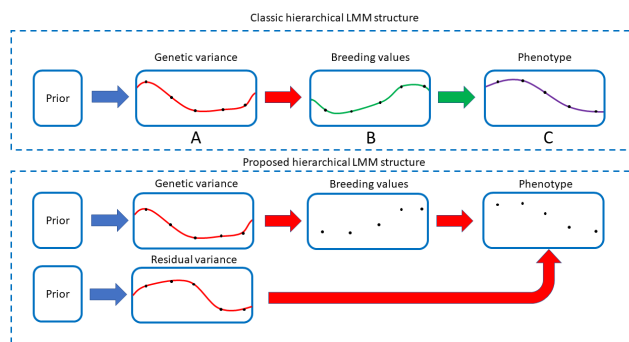


Fig. 1. (Top panel) Continuous smoothing over time points can be induced into one of three possible layers in the LMM hierarchy: either in the variance component layer (A), genetic value layer (B), or phenotype layer (C). (Bottom panel) Our suggested LMM structure, where continuous smoothing over time points is induced on the genetic- and residual variance components.

size (Stinchcombe *et al.*, 2012). In all such cases, it is natural to consider and estimate heritability as a dynamic function. In particular, functional variation in heritability is motivated by the fact that environmental changes may affect the environmental variance, and genes that control traits can activate or deactivate (Bryois *et al.*, 2017), which then may affect the genetic variance. In general, this kind of dynamic modelling of biological processes is a fast growing field thanks to modern data collection techniques, see e.g. Moore *et al.* (2013) and Li and Sillanpää (2015). However, the number of statistical methods and associated (easy-to-use and publicly available) software packages to estimate dynamic heritability is still limited. To discuss this further, let us first put available methods into context. For approaches in the multivariate LMM framework we have the following options.

- i.) A simple approach: Considers all time points as dependent traits and estimates their trait-specific genetic variances. This multi-trait model estimates variances jointly but does not apply any smoothing over time points (Henderson and Quaas, 1976; Lee and van der Werf, 2016).
- ii.) *Smoothing at phenotype level* (Figure 1C): fits a linear or nonlinear function over phenotypic time points, and then estimates latent-trait heritability influencing each parameter of that function using univariate or multivariate LMMs (e.g., Canaza-Cayo *et al.*, 2015).
- iii.) *Smoothing at breeding value level* (Figure 1B): fits a linear or non-linear function over time in genomic breeding values (and possibly residual effects, i.e. permanent environmental effects). A common approach referred to as random regression (Schaeffer, 2016) (<http://animalbiosciences.uoguelph.ca/%7Elrs/BOOKS/rmbook.pdf>), (Campbell *et al.*, 2018), also requires estimation of the residual covariance matrix if the permanent environmental effects are not smoothed.
- iv.) *Smoothing at variance component level* (Figure 1A): If the residual variance is also smoothed, the residual covariance matrix can be left out of the model. Such an example is given by a spline-based method He *et al.* (2016, 2017) which is restricted to only twin data.

In comparison, for MLA approaches we have the two possibilities.

- v.) *Smoothing at phenotype level*: fits a linear or nonlinear function over phenotypes, and then estimates latent-trait heritability for trend parameters (e.g., intercepts, slopes), by either using univariate or multivariate methods (see eg. Gee *et al.*, 2003; Heuven and Janss, 2010; Sillanpää *et al.*, 2012; Li *et al.*, 2014).

- vi.) *Smoothing at quantitative trait locus (QTL) effect level*: fits a linear or non-linear function over time to QTL effects. In particular, these methods have been developed for QTL mapping of function-valued traits, but can also be applied for SNP-heritability estimation. Examples are varying-coefficient models such as Li and Sillanpää (2013) and Vanhatalo *et al.* (2019). Considering the residual covariance structure in these models is recommended.

Dynamic modelling of biological phenomena can be beneficial for a number of reasons. For instance, by improving the precision of the estimation (Li and Sillanpää, 2013). This is due to the simple fact that, as with any statistical modelling, a larger data set leads to increased precision, and naturally the amount of data increases with the number of measurement points. More importantly, if the nature of a phenomenon is dynamic, it is only rational to model the phenomenon dynamically and exploiting information content over time. Finally, there is a heated debate around heritability estimation as to why the current analysis methods are leading to a noticeable gap (known as missing heritability) between heritability estimates from SNP and pedigree / twin data (see e.g., Eichler *et al.*, 2010; Gibson, 2012; Young, 2019). One important contributor to this is the time-dependent nature of heritability, which our method aims to address.

One way to induce smoothing in the model (to model dynamic phenomena) is given by Gaussian processes. These are random processes whose degree of smoothness can be controlled by varying a set of parameters. In particular, they are appealing because of their analytical properties: with certain conjugacy structure the end result is given as an easy-to-calculate formula (Rasmussen and Williams, 2006). Furthermore, Gaussian processes have been successfully utilised in genetics before, for example by Vanhatalo *et al.* (2019), where associations between molecular markers and function-valued traits were studied. Also, the idea of using Gaussian processes in modelling dynamic biological processes is not new. For example, Pletcher and Geyer (1999) and Jaffrézic and Pletcher (2000) suggested that the genetic breeding values and environmental terms of the phenotype in LMM could be viewed as Gaussian processes. They also considered multivariate phenotypes by modelling the covariance function of a process through a parametric representation, which reduces the number of parameters in the model. In fact, this approach differs from ours, because we view the variance components of the LMM as Gaussian processes, which is similar to He *et al.* (2016, 2017), where variance processes were modelled with splines.

In this paper we specifically consider the Bayesian framework of statistical modelling. This means that the statistical inference is not only based on the measurement data and the statistical model, but also on prior assumptions and information about the subject (Gelman *et al.*, 2013). For example, a common assumption in many research fields about dynamic processes is given by their smoothness, i.e. values at neighbouring time points are expected to be closer than at time points further apart. To estimate the parameters in Bayesian models, it is common to use Markov Chain Monte Carlo (MCMC) simulation methods (Robert and Casella, 2009), but since Gaussian process models usually have a high number of highly correlated parameters, efficient MCMC sampling is difficult. Thus, in addition to the traditional Metropolis–Hastings algorithm, we use the state-of-the-art method of elliptical slice sampling (Murray *et al.*, 2010), which has shown to perform well in Gaussian process models.

In addition to dynamic heritability estimation, the aim of this paper is to illustrate the difference in the uncertainty of the estimates of joint and independent modelling of dependent traits. If a trait is measured longitudinally over time, the different measurements can be thought as individual traits. If there is a dependence between traits, it makes sense to model them jointly. This dependence can be viewed as increased sample size because the value of one trait affects the value of the other.

This paper is organised as follows: In section 2 we present our proposed model for dynamic heritability estimation based on Gaussian processes. In addition, we introduce a two-stage method that is used as a baseline. In section 3, we evaluate our method with two simulated and two real datasets and compare to a random regression model implemented with MTG2 (Lee and van der Werf, 2016) and BLUPF90 (Miszta *et al.*, 2002) softwares and ACEt R-package for estimating dynamic heritability (He *et al.*, 2017). In section 4 we examine the obtained results, followed by a discussion in section 5.

2 Methods

We will address the dynamic estimation of heritability by two methods, a joint estimation and a separated two-stage approach which serves to illustrate why joint modelling is beneficial. Specifically, the two-stage method first estimates the posterior means of the variance components and their credible intervals at every time point individually and subsequently combines the obtained estimates into smooth curves.

All models in this study are extensions of the basic linear mixed model, defined as (Henderson, 1984)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}, \quad (1)$$

where $\mathbf{y} \in \mathbb{R}^N$ is a measurement vector, \mathbf{X} is a matrix of fixed effects, $\boldsymbol{\beta}$ contains the regression coefficients associated with them and N is the number of individuals. The matrix \mathbf{Z} (in this case $\mathbf{Z} = \mathbf{I}$) connects random effects $\mathbf{u} \in \mathbb{R}^N \sim \mathcal{N}(\mathbf{0}, \sigma_G^2 \mathbf{G})$ with correct individuals. Here, σ_G^2 is the genetic variance and \mathbf{G} is the genomic relationship matrix (defined in detail below). Lastly, $\boldsymbol{\epsilon} \in \mathbb{R}^N \sim \mathcal{N}(\mathbf{0}, \sigma_E^2 \mathbf{I})$ is the error term. Since we do not have any fixed effects, we can express the centered measurement vector as $\mathbf{y}_c = \mathbf{u} + \boldsymbol{\epsilon}$. We note that the overall covariance matrix of \mathbf{y}_c can be written as $\mathbf{K} = \sigma_G^2 \mathbf{G} + \sigma_E^2 \mathbf{I}$. Narrow-sense heritability is then defined as $h^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_E^2)$.

2.1 Genomic relationship matrix

To separate the genetic and environmental variance components, knowledge of the genetic relations between individuals is needed. A genomic relationship matrix can be estimated from molecular marker data. VanRaden (2008) describes the procedure of constructing such a matrix. First, let \mathbf{M} be an $N \times M$ genotype matrix where N is the number of individuals and M is the number of markers. The elements in \mathbf{M} are coded as $-1, 0, 1$ for homozygote, heterozygote and the other homozygote, respectively. Second, let p_i be the allele frequency of the second allele at locus i . Then we construct a matrix \mathbf{P} with the same dimensions as \mathbf{M} and set the i :th column of \mathbf{P} to $2(p_i - 0.5)$. Let $\mathbf{Q} = \mathbf{M} - \mathbf{P}$. Finally, we can construct a genomic relationship matrix as

$$\mathbf{G} = \frac{\mathbf{Q}\mathbf{Q}'}{2 \sum_i p_i (1 - p_i)}. \quad (2)$$

To reduce the estimation error, we used a shrinkage estimated version of this matrix, defined in detail by Endelman and Jannink (2012) and implemented in the R-package `rRBLUP` (Endelman, 2011). We note that the estimated genomic relationship matrix \mathbf{G} is positive-semidefinite and hence might not be invertible. However, for our model formulation this will not be a problem (for more details, see eq. (9)).

2.2 Joint model over time points

To extend the model (1) to the case of longitudinal data, we can use the Kronecker product denoted by \otimes . The extended model defines a

probability distribution for a data vector $\tilde{\mathbf{y}}_c = [\mathbf{y}_c(1) \dots \mathbf{y}_c(T)]'$ that can be written as

$$\tilde{\mathbf{y}}_c = \tilde{\mathbf{u}} + \tilde{\boldsymbol{\epsilon}}, \quad \tilde{\mathbf{y}}_c, \tilde{\mathbf{u}}, \tilde{\boldsymbol{\epsilon}} \in \mathbb{R}^{NT} \quad (3)$$

where $\tilde{\mathbf{u}} = [\mathbf{u}(1) \dots \mathbf{u}(T)]' \sim \mathcal{N}(\mathbf{0}, \text{diag}(\sigma_G^2(1), \dots, \sigma_G^2(T)) \otimes \mathbf{G})$ contains all breeding values and $\tilde{\boldsymbol{\epsilon}} = [\boldsymbol{\epsilon}(1) \dots \boldsymbol{\epsilon}(T)]' \sim \mathcal{N}(\mathbf{0}, \text{diag}(\sigma_E^2(1), \dots, \sigma_E^2(T)) \otimes \mathbf{I})$ contains all error terms from individual time points consecutively and T is the number of time points. The overall covariance matrix for $\tilde{\mathbf{y}}_c$ can be written as $\tilde{\mathbf{K}} = \text{diag}(\sigma_G^2(1), \dots, \sigma_G^2(T)) \otimes \mathbf{G} + \text{diag}(\sigma_E^2(1), \dots, \sigma_E^2(T)) \otimes \mathbf{I}$. The variance component vectors can now be estimated with a Bayesian approach by setting suitable priors for them. It is worth noting that this model structure implies that all covariances across time (within and between individuals) are ignored. For the rapid performance of the algorithm this assumption is crucial. It is in the priors of the variance component vectors where we assume dependence over time (see Figure 1, bottom panel). This is a fundamental difference compared to for example Pletcher and Geyer (1999) who model the dependence on the level of breeding values. The idea behind our assumption is that the dependence of neighbouring trait values induces dependence on the variance level.

The priors are based on assumptions about the qualitative features of the variance component functions, namely their continuity and smoothness. Due to the aforementioned reasons, we assume that the variance values at neighbouring time points are on average closer than at time points further apart. To formulate such assumptions mathematically, one can use Gaussian processes, fully defined by a mean function and a covariance function (Rasmussen and Williams, 2006). We write $\log \sigma_G^2 \sim \mathcal{GP}(0, C_G(t, t'))$ and $\log \sigma_E^2 \sim \mathcal{GP}(0, C_E(t, t'))$, where $C_G(t, t') = \text{cov}(\log \sigma_G^2(t), \log \sigma_G^2(t'))$ and $C_E(t, t') = \text{cov}(\log \sigma_E^2(t), \log \sigma_E^2(t'))$. Logarithms are used to allow negative values for the processes and the zero mean can be achieved approximately by suitable translation of the data. For C_G and C_E , we use the Matérn covariance function, given by

$$C(t, t') = \tau^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{|t - t'|}{\lambda} \right)^\nu K_\nu \left(\frac{|t - t'|}{\lambda} \right), \quad (4)$$

where τ^2 is the magnitude parameter that controls the overall variance of the stochastic process, λ is the length scale that governs how fast the covariance drops with respect to the distance of time points, $\Gamma(\cdot)$ denotes the gamma function and $K_\nu(\cdot)$ the modified Bessel function of the second kind. The value of ν controls the mean-square differentiability of the process, which affects its smoothness. We note that for $\nu = 0.5, 1.5, 2.5, \dots$, there exists a simple form for the covariance function. In this study, we fix $\nu = 1.5$, which means that a sample process is once mean-square differentiable and we obtain

$$C_{1.5}(t, t') = \tau^2 \left(1 + \frac{|t - t'|}{\lambda} \right) \exp \left(-\frac{|t - t'|}{\lambda} \right) \quad (5)$$

(Rasmussen and Williams, 2006). This means the process is rather smooth, but it is also allowed to change rapidly.

A common problem with Gaussian processes is determining values for the hyperparameters τ^2 and λ . In the Bayesian framework one can set up separate priors for each and estimate them together with the other unknowns using MCMC. However, it is known that identifiability issues arise when simultaneously estimating both hyperparameters (Zhang, 2004). Thus, it is common to fix one of them. In our case, we fix $\tau^2 = 1$ and scale the data so that the mean of the overall variance over time points equals 2. This means the mean of the environmental and genetic variance over time points is 1 if they are equal.

Because our Gaussian process model is nonparametric, we have to discretise the time axis and define a grid of points where we estimate the

variance components. For simplicity, we estimate the variance components at the same locations the measurements were taken. This means that the covariance function $C(\cdot, \cdot)$ becomes a covariance matrix \mathbf{C} . Then we have for each element in the matrix that $[C]_{ij} := C(t_i, t_j)$. We note that a covariance matrix defined this way is dense and computationally heavy to operate with. Hence we use a sparse approximation for the inverse of the covariance matrix (Roininen *et al.*, 2014), discussed in more detail in the supplementary material.

The model can be written in hierarchical form as

$$\begin{aligned} \tilde{\mathbf{y}}_c | \sigma_E^2, \sigma_G^2 &\sim \mathcal{N}(\mathbf{0}, \tilde{\mathbf{K}}), \\ \log \sigma_E^2 | \lambda_E &\sim \mathcal{N}(\mathbf{0}, \mathbf{C}_E), \\ \log \sigma_G^2 | \lambda_G &\sim \mathcal{N}(\mathbf{0}, \mathbf{C}_G), \\ \log \lambda_E &\sim \mathcal{N}(\mu_\lambda, \zeta_\lambda^2), \\ \log \lambda_G &\sim \mathcal{N}(\mu_\lambda, \zeta_\lambda^2). \end{aligned} \quad (6)$$

By Bayes' formula, the unnormalised posterior density can be expressed as

$$\begin{aligned} p(\sigma_G^2, \sigma_E^2, \lambda_E, \lambda_G | \tilde{\mathbf{y}}_c) &\propto \\ p_{\mathcal{N}}(\tilde{\mathbf{y}}_c | \mathbf{0}, \tilde{\mathbf{K}}) p_{\mathcal{N}}(\log \sigma_G^2 | \mathbf{0}, \mathbf{C}_G) p_{\mathcal{N}}(\log \sigma_E^2 | \mathbf{0}, \mathbf{C}_E) \\ p_{\mathcal{N}}(\log \lambda_E | \mu_\lambda, \zeta_\lambda^2) p_{\mathcal{N}}(\log \lambda_G | \mu_\lambda, \zeta_\lambda^2), \end{aligned}$$

where $p_{\mathcal{N}}$ refers to the multinormal density.

In practice, with this parameterization we observed poor mixing of the hyperparameter chains. Hence, we employed whitening (Murray and Adams, 2010; Yu and Meng, 2011; Monterrubio-Gómez *et al.*, 2019), which breaks the dependencies of the variance component processes and corresponding hyperparameters under the prior. We first note that the logarithmized variance component processes can be expressed as $\log \sigma_E^2 = \mathbf{C}_E^{\frac{1}{2}} \boldsymbol{\eta}_E$ and $\log \sigma_G^2 = \mathbf{C}_G^{\frac{1}{2}} \boldsymbol{\eta}_G$, where $\boldsymbol{\eta}_E$ and $\boldsymbol{\eta}_G$ are both standard multivariate normally distributed. The matrix square roots are obtained straightforwardly through the approximations (see supplementary material). Instead of $\log \sigma_E^2$ and $\log \sigma_G^2$, we now sample $\boldsymbol{\eta}_E$ and $\boldsymbol{\eta}_G$. The reparametrisation corresponds to the following posterior density

$$\begin{aligned} p(\boldsymbol{\eta}_G, \boldsymbol{\eta}_E, \lambda_E, \lambda_G | \tilde{\mathbf{y}}_c) &\propto p_{\mathcal{N}}(\tilde{\mathbf{y}}_c | \mathbf{0}, \tilde{\mathbf{K}}) p_{\mathcal{N}}(\boldsymbol{\eta}_G | \mathbf{0}, \mathbf{I}) p_{\mathcal{N}}(\boldsymbol{\eta}_E | \mathbf{0}, \mathbf{I}) \\ &\quad p_{\mathcal{N}}(\log \lambda_E | \mu_\lambda, \zeta_\lambda^2) p_{\mathcal{N}}(\log \lambda_G | \mu_\lambda, \zeta_\lambda^2), \end{aligned}$$

where

$$\tilde{\mathbf{K}} = \text{diag} \left(\exp \left\{ \mathbf{C}_G^{\frac{1}{2}} \boldsymbol{\eta}_G \right\} \right) \otimes \mathbf{G} + \text{diag} \left(\exp \left\{ \mathbf{C}_E^{\frac{1}{2}} \boldsymbol{\eta}_E \right\} \right) \otimes \mathbf{I}.$$

The hyperparameters μ_λ and ζ_λ^2 are set by following Monterrubio-Gómez *et al.* (2019). In particular, it is based on the idea that the length scale is identifiable between the smallest and largest distance between two time points, say $[a, b]$. Hence we want to place most of its prior probability mass in that interval. By using the quantile function of a standard normal distribution, we can assign approximately 95% of the prior mass between the interval by solving the following system of equations:

$$\begin{aligned} \mu_\lambda - 1.96\zeta_\lambda &= \log a \\ \mu_\lambda + 1.96\zeta_\lambda &= \log b. \end{aligned} \quad (7)$$

We emphasize that this model specification leaves no tuning parameters to fix prior to estimation which eliminates the need for preliminary analyses and consequently saves plenty of time.

2.2.1 Parameter estimation in the joint model

To generate dependent MCMC samples from the posterior distribution of parameters $\boldsymbol{\eta}_G$ and $\boldsymbol{\eta}_E$ we use the elliptical slice sampling method (Murray *et al.*, 2010). It is a rejection-free sampling algorithm and we noticed that it does perform better for high dimensional data (more than 100 time points) than the block-update of Metropolis–Hastings (MH) algorithm. The problem with elliptical slice sampling is that the likelihood must be evaluated multiple times in a single MCMC-iteration, making it quite a bit slower than MH. For the length scales we use MH with Gaussian random walk proposals, i.e. the proposal value is sampled from the distribution $\mathcal{N}(\theta^{(t)}, \sigma_{\text{RW}}^2)$, where $\theta^{(t)}$ is the current value of the parameter. The variance is adapted following section 3 in Roberts and Rosenthal (2009) to achieve an acceptance rate of 0.44 which is considered optimal in certain settings. The four parameter subsets $(\boldsymbol{\eta}_E, \boldsymbol{\eta}_G, \log \lambda_E, \log \lambda_G)$ are sampled in alternatingly while keeping the other values fixed.

The computationally most intensive part of the algorithm is the evaluation of the likelihood function, due to the inversion and determinant computation of an $NT \times NT$ covariance matrix. Fortunately, the form of the matrix allows us to utilise some basic linear algebra to make the computations feasible. We start by examining the structure of matrix $\tilde{\mathbf{K}}$ and note the block diagonal structure where the blocks consist of time point specific covariance matrices:

$$\tilde{\mathbf{K}} = \begin{bmatrix} \sigma_G^2(1)\mathbf{G} + \sigma_E^2(1)\mathbf{I} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_G^2(T)\mathbf{G} + \sigma_E^2(T)\mathbf{I} \end{bmatrix}.$$

This means we can express the overall log-likelihood function of the parameters as a sum of log-likelihood functions of each time point

$$\log p(\tilde{\mathbf{y}}_c | \mathbf{0}, \tilde{\mathbf{K}}) = \sum_{t=1}^T \log p(\mathbf{y}_c(t) | \mathbf{0}, \mathbf{K}(t)). \quad (8)$$

We can now use eigen decomposition to decorrelate the measurements using a linear transformation to speed up the computations. The covariance matrix of the measurements $\mathbf{y}(t)$ at time t can be decomposed as $\mathbf{U}\mathbf{D}(t)\mathbf{U}'$, where \mathbf{U} is the eigenvector matrix of \mathbf{G} and $\mathbf{D}(t)$ is a diagonal matrix with $[D(t)]_{nn} = \sigma_G^2(t)\xi_n + \sigma_E^2(t)$ and ξ_n being the eigenvalues of \mathbf{G} . By the orthogonality of \mathbf{U} , we have that $\text{var}(\mathbf{U}'\mathbf{y}(t)) = \mathbf{U}'\mathbf{U}\mathbf{D}(t)\mathbf{U}'\mathbf{U} = \mathbf{D}(t)$, implicating that the elements in the transformed measurement vector $\mathbf{U}'\mathbf{y}(t)$ are independent of each other. This reduces the log-likelihood calculation into a sum. Most importantly, since \mathbf{U} does not depend on the variance components, the transformation has to be done only once. Setting $\mathbf{z}(t) = \mathbf{U}'\mathbf{y}(t)$, the resulting log-likelihood can be written as

$$\log p(\tilde{\mathbf{y}} | \mathbf{0}, \tilde{\mathbf{K}}) = \sum_{t=1}^T \sum_{n=1}^N \log p(z(t)_n | 0, \sigma_G^2(t)\xi_n + \sigma_E^2(t)). \quad (9)$$

The pseudocode for the algorithm can be found in the supplementary material (algorithm 1).

2.3 Two-stage method

To illustrate the benefits of the joint model, we compare it to a baseline approach. The method consists of two stages: (i.) the estimation of posterior means of the variance components and their 95% credible intervals at each time point separately and (ii.) combining individual estimates together by smoothing the obtained estimates from stage one. The estimation in stage one is based on the same model as in the joint

estimation but for one time point

$$\mathbf{y}_c(t) = \mathbf{u}(t) + \boldsymbol{\epsilon}(t), \quad (10)$$

where \mathbf{y}_c is the centered measurement vector at time t , $\mathbf{u}(t) \sim \mathcal{N}(\mathbf{0}, \sigma_G^2(t)\mathbf{G})$ and $\boldsymbol{\epsilon}(t) \sim \mathcal{N}(0, \sigma_E^2(t)\mathbf{I})$. As in the joint method, the data is scaled such that the overall variance over time points is 2. The covariance matrix of $\mathbf{y}_c(t)$ is $\mathbf{K}(t) = \sigma_G^2(t)\mathbf{G} + \sigma_E^2(t)\mathbf{I}$ and the model can be written as

$$\begin{aligned} \mathbf{y}_c(t) &\sim \mathcal{N}(\mathbf{0}, \mathbf{K}(t)), \\ \log \sigma_G^2(t) &\sim \mathcal{N}(0, 1), \\ \log \sigma_E^2(t) &\sim \mathcal{N}(0, 1). \end{aligned} \quad (11)$$

We note that this is a special case of the joint model presented earlier, see eq. (3). Posterior means and 95% credible intervals of the variance components at each time point are obtained from this analysis used in the next stage.

Smoothing of the variance component curves over time is based on the model

$$\mathbf{y}_{\sigma^2} = \mathbf{f}_{\sigma^2} + \boldsymbol{\epsilon}_{\sigma^2}, \quad (12)$$

where \mathbf{y}_{σ^2} contains either logarithmized posterior means, lower 95% credible interval limits or upper 95% credible interval limits estimated in stage one, $f_{\sigma^2} \sim \mathcal{GP}(0, C_\lambda(t, t'))$ is a smooth process and $\boldsymbol{\epsilon}_{\sigma^2} \sim \mathcal{N}(\mathbf{0}, \gamma^2\mathbf{I})$ is an error term. $C_\lambda(\cdot, \cdot)$ is again the Matérn covariance function, defined in eq. (5). Here \mathbf{y}_{σ^2} is scaled to have mean zero and variance one. After discretisation, the model can be written as follows:

$$\begin{aligned} \mathbf{y}_{\sigma^2} &\sim \mathcal{N}(\mathbf{0}, \mathbf{C}_\lambda + \gamma^2\mathbf{I}), \\ \log \gamma^2 &\sim \mathcal{N}(0, 10000), \\ \log \lambda &\sim \mathcal{N}(\mu_\lambda, \zeta_\lambda^2), \end{aligned} \quad (13)$$

where γ^2 and λ are assumed to be *a priori* independent. The variance parameter in the error term is given an uninformative prior and the parameters μ_λ and ζ_λ^2 are fixed similarly as in joint estimation method following Monterrubio-Gómez *et al.* (2019). The process f_{σ^2} is analytically integrated out of the model, but we can reconstruct it at the measurement points after the estimation of the hyperparameters (Rasmussen and Williams, 2006) by noting that

$$\begin{aligned} \mathbf{f}_{\sigma^2} | \mathbf{y}_{\sigma^2}, \lambda, \gamma^2 &\sim \\ \mathcal{N}(\mathbf{C}_\lambda(\mathbf{C}_\lambda + \gamma^2\mathbf{I})^{-1}\mathbf{y}_{\sigma^2}, \mathbf{C}_\lambda - \mathbf{C}_\lambda(\mathbf{C}_\lambda + \gamma^2\mathbf{I})^{-1}\mathbf{C}_\lambda). \end{aligned}$$

2.3.1 Parameter estimation in the two-stage method

The parameters $\log \sigma_G^2(t)$ and $\log \sigma_E^2(t)$ in stage one are estimated using a random walk MH-algorithm with the same adaptation as in the joint estimation method. They are sampled at MCMC iteration $i + 1$ from the distributions

$\mathcal{N}(\log \sigma_G^2(t)^{(i)}, s_{\sigma_G^2}(t)^{(i)})$ and $\mathcal{N}(\log \sigma_E^2(t)^{(i)}, s_{\sigma_E^2}(t)^{(i)})$. The proposals are accepted by MH (a single parameter at a time) conditionally on the other parameter fixed to its latest value.

The parameters $\log \gamma^2$ and $\log \lambda$ of the smoothing model in phase two are estimated similarly as in stage one with the same adaptation. They are sampled at iteration $i + 1$ from the distributions $\mathcal{N}(\log \gamma^2^{(i)}, s_{\gamma^2}^{(i)})$ and $\mathcal{N}(\log \lambda^{(i)}, s_\lambda^{(i)})$. The proposals are accepted by MH (a single parameter at a time) conditionally on the other parameter fixed to its latest value.

The pseudocode for both stages can be found in the supplementary material (algorithms 2 and 3).

2.4 Computational considerations

Since the log-likelihood in the joint estimation method can be expressed as a sum (eq. 9), its computation can be parallelized. We parallelized all log-likelihood calculations, gaining additional speedups. The computation times were between 24 minutes for the mouse activity data and 31 minutes for the *Arabidopsis* data with 300 000 MCMC-iterations. The workstation used for the simulations had an AMD Ryzen Threadripper 2950X 3.5 Ghz processor with 16 cores and 32 GB of RAM. The method is implemented with C++ integrated with R using the Rcpp library (Eddelbuettel and François, 2011) and Eigen (Guennebaud *et al.*, 2010) (<http://eigen.tuxfamily.org>) for linear algebra. The program is available at <https://github.com/aarjas/dynBGP>.

3 Example analyses

To test the methods described above, we used four different datasets with a large number of time points. Two were simulated and the two others were real datasets from <https://phenome.jax.org/centers/QTLA>.

3.1 Simulated dataset

The simulated data consists of measurements from $N = 1000$ individuals from $T = 50$ time points. First, a relationship matrix was created by generating an $N \times N$ matrix \mathbf{S} with independent standard normally distributed elements. The relationship matrix was then computed as $\mathbf{G} = \mathbf{S}\mathbf{S}^T/N + 0.1\mathbf{I}$. A small number was added to the diagonal to make the matrix positive definite. To simulate realistic data, the longitudinal dependencies need to be taken into account as well. This was done by computing a $T \times T$ Gaussian process matrix \mathbf{C} where the row i and column j intersection was set to $[C]_{ij} = \left(1 + \frac{\sqrt{5}|t_i - t_j|}{(50/3)} + \frac{5(t_i - t_j)^2}{3 \cdot (50/3)^2}\right) \exp\left(-\frac{\sqrt{5}|t_i - t_j|}{(50/3)}\right)$. The genetic and environmental components of the data were simulated from the distributions $\mathcal{N}(\mathbf{0}, \mathbf{C} \otimes \mathbf{G})$ and $\mathcal{N}(\mathbf{0}, \mathbf{C} \otimes \mathbf{I})$, respectively. Finally, the components were scaled with the corresponding variances at each time point and summed. A correct relationship matrix was assumed known in the analysis stage. To demonstrate scalability of the approach, we created larger datasets with up to 50 000 individuals and 1000 time points, see supplementary material (Fig. S5).

3.2 *Arabidopsis thaliana* dataset

The second dataset contains *Arabidopsis thaliana* root tip angle measurements (Moore *et al.*, 2013). The population consists of $N = 162$ recombinant inbred lines of *Arabidopsis* seeds with 234 markers. The seeds were placed on Petri dishes that were held in front of a camera and rotated 90° so that the roots grew parallel to the ground. The root tip angle was measured every two minutes for eight hours resulting in $T = 239$ time points. Moore *et al.* (2013) describe the measurement process in more detail. We used a version of the data where the phenotype values were averages of multiple individuals representing the same line.

3.3 Mouse activity dataset

The third dataset consists of mouse activity measurements (Xiong *et al.*, 2011). The activity of $N = 89$ mice was monitored over a period of 12 days. One day was divided into 6 minute intervals and an active state probability of a given mouse in each interval was calculated based on the 12 days of data. This results in $T = 222$ time points. Since the outcome is a probability, normality assumption is not realistic. Hence we transformed the phenotypic data using logit-transformation following the procedure described by Li and Sillanpää (2013). The data also has measurements of 251 markers.

3.4 Comparison with ACET

We also compared our joint estimation method with the freely available ACET R-package for estimating dynamic heritability (He *et al.*, 2016, 2017). The method is restricted only for twin data and it uses splines in defining the different dynamic variance components. The ACET model also includes a common environmental effect as a third variance component, which our model lacks. For uncertainty quantification, ACET uses delta-method or bootstrap. For the method comparison we used a simulated twin dataset that comes with the ACET R-package. It consists of 100 twin pairs, half of who are monozygotic and half of who are dizygotic. The $N = 200$ individuals have $T = 50$ equispaced measurements of an artificial trait over 50 years. In twin models, the covariances in the relationship matrix between a monozygotic twin pair and a dizygotic twin pair are assumed to be 1 and 0.5, respectively, while the diagonal values are all 1. Between-pair covariance is assumed to be 0.

3.5 Completing missing genotype data by imputation

There were no missing phenotype values in any of the datasets. The mouse activity dataset had 0.3% and the *Arabidopsis* dataset 1.4% missing marker values. All missing values were imputed once before calculating the genomic relationship matrix with the mean value of the given marker over individuals.

4 Results

The results from the Bayesian analysis of simulated data are presented in Figure 2 for both methods. The posterior estimated variance component functions generally follow the real functions that were used to generate the data well. Perhaps the most interesting aspect about the results is the difference in the uncertainty of the estimates between the methods. The 95% credible intervals provided by the joint model are far narrower than those of the two-stage method. This is because the joint model makes use of the whole dataset, while the two-stage method uses the data of each time point independently.

The estimated posterior mean curves and their 95% credible intervals for variance components and SNP-heritability calculated from the *Arabidopsis* seed data are presented for the joint model and two-stage method in Figure 3. Again, one can clearly see much wider 95% credible intervals surrounding the posterior mean curves obtained from the two-stage method. Moore *et al.* (2013) have also estimated the dynamic heritability from this dataset using ANOVA. They calculated the variances within and between genetically distinct lines separately at every point in time. The curve has the same shape but the actual values are somewhat smaller than ours, peaking at 0.25, while our estimate peaks at about 0.4. This offset can be explained by different averaging of used data and in particular, their estimation lacks uncertainty estimates. Vanhatalo *et al.* (2019) have analyzed the same data as well and are likely using the same averaging as we do. We note, that their obtained heritability estimates coincide very well with ours. However, they also lack uncertainty estimates. Interestingly, similar results were not expected, since the two analysis models differ significantly in structure and especially in terms of assumptions on the underlying genetics.

The posterior estimates from the mouse activity data are presented in Figure 4. This kind of data seems to be challenging for the Gaussian process models. This is mainly due to the fact that here the variance processes have very rough features along with smooth areas. Thus, the amount of smoothing needed is very different at different time points. Especially the heritability process estimated with the two-stage method does not perform so well and produces highly oscillating features. The same data was analysed in Vanhatalo *et al.* (2019) as well. In this case, their

heritability estimates look quite different to ours. In particular, our methods have smoothed most rough edges, whereas their estimate has preserved them. Nevertheless, the overall shape is similar. These differences here are likely the consequence of using very different models.

We also compared the joint model with a random regression model (RRM) (implemented in for example BLUPF90 (Misztal *et al.*, 2002), MCMCglmm (Hadfield, 2010) and MTG2 (Lee and van der Werf, 2016)) which is a well established method for analysing dynamic biological phenomena. We simulated ten different datasets with the method from section 3.1 and fitted a Legendre polynomial of degree five for both the genomic breeding values and permanent environmental effects. The model assumes heterogenous residual variances over time. A more precise model definition can be found in the supplementary material. We computed the mean squared error (MSE) of the estimated genetic and environmental variances as well as heritabilities, where $MSE = (1/T) \sum_{t=1}^T (f(t) - \hat{f}(t))^2$, with $f(t)$ being the ground truth at time t and $\hat{f}(t)$ the estimate at time t . We used posterior mean as the point estimate for the joint method. The computed errors can be found in the supplementary material (Table S1). This analysis was performed using AI-REML method in the software MTG2 (Lee and van der Werf, 2016) and it took on average 14 minutes. The MSE of the joint method was on average 17% smaller than of the RRM. We note that the MSE varies with the chosen degree of the Legendre polynomials and for this computation we chose the best performing combinations with a reasonable computation time. Additionally, we fitted a Bayesian version of the same model using BLUPF90 family of programs (Misztal *et al.*, 2002) and GIBBS2F90 in particular. This allowed us to quantify the uncertainty in the variance components estimated with the RRM. The results of the analysis and further details can be found in the supplementary material (Fig. S7). The 95% credible intervals obtained from the RRM are slightly wider than of the joint model. This might be due to the differences in model structures (cf. Fig. 1). The computation time for the Bayesian RRM for the simulated dataset was over 24 hours while for the joint method it was less than 30 minutes with 300 000 MCMC-iterations in both. Furthermore, we would like to mention that the RRM implemented in both MTG2 and GIBBS2F90 did not converge on our real dataset examples where the number of individuals is small but the number of time points high. Yet, the joint method manages to do so, i.e. the MCMC sampler converges. It is not completely clear to us why this is the case. One possibility might be the difference in assumptions: In RRM the assumptions about the parametric shape of the breeding values and environmental effects induce some shape for the variance components. In the joint method, the prior assumptions concern only the variances. This direct modelling strategy might simplify the estimation process.

Finally, the comparison of our joint model with the ACET model for analyzing twin data is presented in the supplementary material (Fig. S4). The results obtained by both methods are consistent, even though ACET considers also common environment in its model. In our method, the genetic variance component seems to have absorbed the common environmental variation, resulting in slightly higher heritability estimate than the one given by ACET. The uncertainty estimates differ a bit between the methods, which is likely due to the common environmental component. Moreover, obtained uncertainty limits are not fully comparable because the frequentist ACET method uses delta-method and our Bayesian method uses MCMC sampling for generating the limits.

Performance of the MCMC-algorithm for the joint model can be evaluated from traceplots of the parameters (supplementary material, Fig. S1–S3). All of the parameter chains seem to have converged, although there is some oscillation on the variance component parameters with the lowest effective sample size.

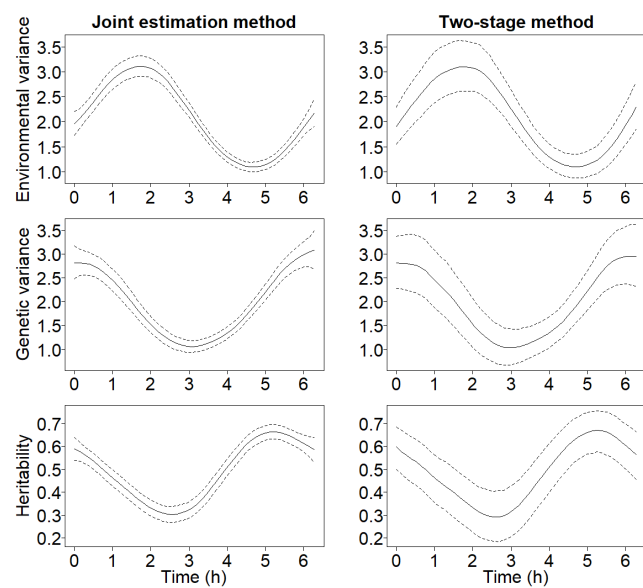


Fig. 2. The dynamic variance components and SNP-heritability estimated from the simulated dataset with the two different Bayesian methods. The posterior mean curves are drawn with solid lines and 95% credible intervals with dashed lines.

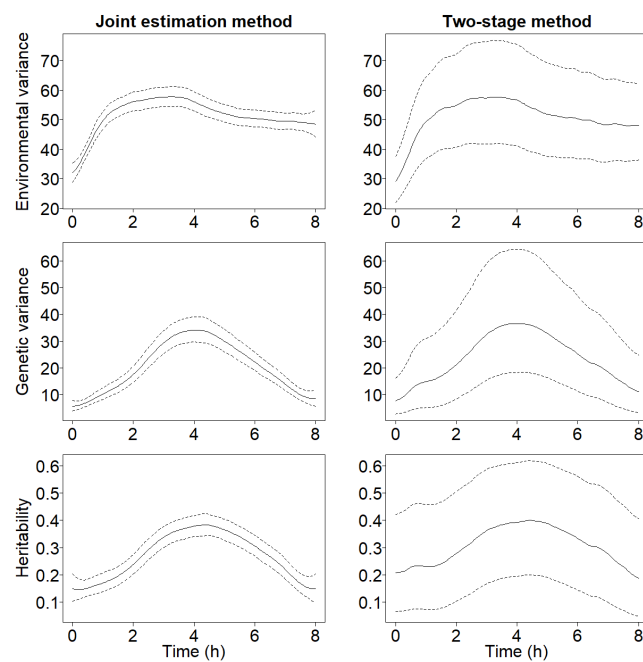


Fig. 3. The dynamic variance components and SNP-heritability estimated from the *Arabidopsis thaliana* dataset with the two different Bayesian methods. Posterior mean curves are drawn with solid lines and 95% credible intervals with dashed lines.

5 Discussion and conclusions

We presented a new Bayesian method for estimating dynamic narrow-sense heritability, based on linear mixed models and Gaussian process priors. The method uses data from all time points at once, making it possible for the time points to 'borrow strength' from one another through the prior covariance structure. This property makes the resulting posterior distributions narrower compared to the second method where the variance components are estimated separately for each time point and smoothed

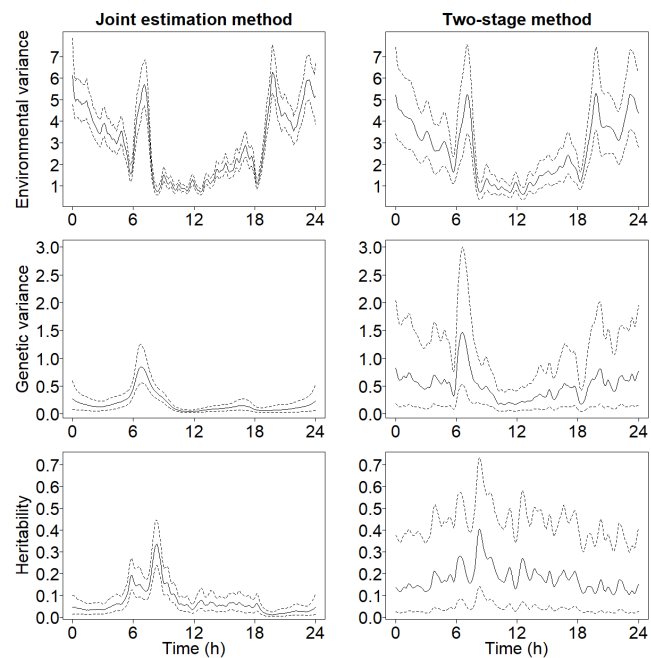


Fig. 4. The dynamic variance components and SNP-heritability estimated from the mouse activity dataset with the two different Bayesian methods. Posterior mean curves are drawn with solid lines and 95% credible intervals with dashed lines.

afterwards. Another benefit of the method presented due to nonparametric smoothing using Gaussian processes is that it can handle very general functional shapes.

The presented estimation method bears some similarities with random regression as both are based on linear mixed models, but there are essential differences that we would like to point out. Most notably, in our proposed method the smoothing is based on priors that are set for the variance component vectors, while in random regression the smoothness of the variances is induced through the assumptions made about the functional shape of the breeding values and environmental effects. Additionally, our model assumes independence of different traits (measurements made at different points in time) and ties them together with the priors. In multivariate linear mixed models, traits are assumed to be dependent and random regression attempts to reduce the size of the estimated covariance matrix by reparametrisation.

A benefit of our proposed method is the ability to quantify the uncertainty in the variance component estimates. This is also possible in Bayesian RRM using MCMC but in frequentistic RRM additional steps such as application of delta method or bootstrap are needed to produce the respective confidence intervals. In addition, our method is completely free of tuning whereas in random regression the degree of the polynomial or alternatively knot points in splines must be chosen prior to estimation. In practice, we noticed that the estimation in MTG2 is really fast for a few traits but it slows down quickly as the number of traits increases. In fact, trying to run the algorithm with 1000 individuals and 100 traits causes an insufficient virtual memory error. In contrast, our algorithm still worked well with 50 000 individuals and 100 time points within a reasonable time (approximately 20 hours). Additionally, high number of traits is not an issue either as demonstrated in the supplementary material (Fig. S5). We like to emphasize that the computation times grow almost linearly and hence our algorithm exhibits excellent scalability. Based on our results, we believe that random regression performs best when there is a high number of individuals and low number of time points which was not the

case in our real data examples. In MTG2, the time complexity is cubic with respect to the number of time points (Lee and van der Werf, 2016), and we believe this is also the case with other RRM implementations. It is also noted by Schaeffer (2016) that Legendre polynomials might cause artefacts near the boundaries of the covariate domain and hence Gaussian processes are favoured. A limitation of our method is the inability to model gene-covariate interactions with possibly only one measurement per individual (cf. Ni *et al.*, 2018; Moore *et al.*, 2019). Theoretically, to perform such analysis, the covariate would have to be split up into discrete groups with each group containing suitably large population (cf. Robinson *et al.*, 2017). This would also result in each group having their own relationship matrix. This is not supported by the algorithm at the moment, however.

Furthermore, the method presented here can also be extended for further analysis. For example, given the posterior mean estimates $\hat{\sigma}_G^2(t)$ and $\hat{\sigma}_E^2(t)$ at time t , the conditional distribution of genomic breeding values $\mathbf{u}(t)$ is $\mathcal{N}(\boldsymbol{\mu}_u(t), \boldsymbol{\Sigma}_u(t))$ (Rasmussen and Williams, 2006), where

$$\begin{aligned}\boldsymbol{\mu}_u(t) &= \hat{\sigma}_G^2(t)\mathbf{G}(\hat{\sigma}_G^2(t)\mathbf{G} + \hat{\sigma}_E^2(t)\mathbf{I})^{-1}\mathbf{y}(t) \text{ and} \\ \boldsymbol{\Sigma}_u(t) &= \hat{\sigma}_G^2(t)\mathbf{G} - \hat{\sigma}_G^2(t)\mathbf{G}(\hat{\sigma}_G^2(t)\mathbf{G} + \hat{\sigma}_E^2(t)\mathbf{I})^{-1}\hat{\sigma}_G^2(t)\mathbf{G}.\end{aligned}\quad (14)$$

The posterior means of genomic breeding values are not smooth functions over time (cf. Campbell *et al.*, 2018). After computing the breeding values, the SNP effects $\mathbf{m}(t)$ can be estimated for association mapping purposes by the back-transformation formula $\hat{\mathbf{m}}(t) = \mathbf{R}'\mathbf{G}^{-1}\boldsymbol{\mu}_u(t)$, where $[R]_{ij} = ([M]_{ij} + 1 - 2p_j)(\sqrt{2p_j(1-p_j)})^{-1}$, p_j is the allele frequency of the other allele in marker j and \mathbf{M} defined in section 2.1 (Bernal Rubio *et al.*, 2016). This kind of longitudinal analysis was recently performed by Campbell *et al.* (2019) who first estimated the breeding values with a random regression model and then used the back-transformation to solve for the SNP effects. Results were compared to an alternative single time point analysis and the authors found that the dynamic analysis recovered more significant associations. The presented model can also be extended by adding fixed environmental effects that affect the phenotype values. This could further reduce the estimation error. Another strategy is to apply a two-stage pre-correction similar to He *et al.* (2016). This means that a linear regression model is first fitted to estimate the fixed effect coefficients and the residuals of that model are then used as phenotype values to estimate the variance components in our LMM.

Further extension would be to consider non-Gaussian longitudinal phenotypes, for example binary or count data. In case of binary data, an extra latent trait layer can be added to the model (Albert and Chib, 1993; Felsenstein, 2005; Kärkkäinen and Sillanpää, 2013). In latent trait modelling, the binary phenotype can be modelled by considering an underlying hypothetical normally distributed latent trait variable which gives rise to the binary trait at the observed layer. If the latent trait variable is smaller than the pre-determined threshold, binary trait at the observed layer obtains the value zero and otherwise it obtains the value one.

One future extension is also to consider multiple longitudinal quantitative traits simultaneously. Some models and methods have been presented for this purpose – to explain the variation of more than one trait simultaneously over time (Sung *et al.*, 2016; Oliveira *et al.*, 2019).

To conclude, we presented a new tuning-free method for estimating dynamic heritability using a Bayesian linear mixed model and GP priors. To estimate the parameters in the model, we use MCMC which makes the uncertainty quantification straightforward. Our results clearly illustrate that joint modelling of the data of all time points reduces the uncertainty in the estimates compared to independent modelling.

Acknowledgements

We are grateful to the editors and three anonymous referees for their constructive comments that helped us to substantially improve the presentation of this article.

Funding

This work was partially supported by the Academy of Finland Profi 5 funding for mathematics and AI: data insight for high dimensional dynamics and the Academy of Finland (Project 312123, Finnish Centre of Excellence in Inverse Modelling and Imaging, 2018–2025). AH acknowledges support from EPSRC grants EP/N032055/1 and EP/M020533/1.

References

- Albert, J. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, **88**, 669–679.
- Bernal Rubio, Y. L. *et al.* (2016). Meta-analysis of genome-wide association from genomic prediction models. *Animal Genetics*, **47**(1), 36–48.
- Bryois, J. *et al.* (2017). Time-dependent genetic effects on gene expression implicate aging processes. *Genome Research*, **27**, 542–552.
- Campbell, M. *et al.* (2018). Utilizing random regression models for genomic prediction of a longitudinal trait derived from high-throughput phenotyping. *Plant Direct*, **2**(9), e00080.
- Campbell, M. *et al.* (2019). Leveraging breeding values obtained from random regression models for genetic inference of longitudinal traits. *The Plant Genome*, **12**(2), 180075.
- Canaza-Cayo, A. *et al.* (2015). Modelling of growth curves and estimation of genetic parameters for growth curve parameters in peruvian young llamas (lama glama). *Small Ruminant Research*, **130**, 81–89.
- Eddelbuettel, D. and François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, **40**(8), 1–18.
- Eichler, E. *et al.* (2010). Viewpoint: Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics*, **11**, 446–450.
- Endelman, J. and Jannink, J.-L. (2012). Shrinkage estimation of the realized relationship matrix. *G3 (Bethesda)*, **2**(11), 1405–1413.
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, **4**, 250–255.
- Felsenstein, J. (2005). Using the quantitative genetic threshold model for inferences between and within species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**(1459), 1427–1434.
- Gee, C. *et al.* (2003). Segregation and linkage analysis for longitudinal measurements of a quantitative trait. *BMC genetics*, **4 Suppl 1**, S21.
- Gelman, A. *et al.* (2013). *Bayesian Data Analysis*. CRC Press LLC, Boca Raton.
- Gibson, G. (2012). Rare and common variants: Twenty arguments. *Nature Reviews Genetics*, **13**, 135–145.
- Guennebaud, G. *et al.* (2010). Eigen v3.
- Hadfield, J. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**(2), 1–22.
- He, L. *et al.* (2016). Estimating modifying effect of age on genetic and environmental variance components in twin models. *Genetics*, **202**, 1313–1328.
- He, L. *et al.* (2017). ACEt: An R package for estimating dynamic heritability and comparing twin models. *Behavior Genetics*, **47**,

- 620–641.
- Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. University of Guelph, Guelph.
- Henderson, C. R. and Quaas, R. L. (1976). Multiple Trait Evaluation Using Relatives' Records. *Journal of Animal Science*, **43**(6), 1188–1197.
- Heuven, H. C. M. and Janss, L. L. G. (2010). Bayesian multi-QTL mapping for growth curve parameters. *BMC Proceedings*, **4**(1), S12.
- Jaffrézic, F. and Pletcher, S. D. (2000). Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics*, **156**(2), 913–922.
- Kärkkäinen, H. P. and Sillanpää, M. J. (2013). Fast genomic predictions via Bayesian G-BLUP and multilocus models of threshold traits including censored Gaussian data. *G3 (Bethesda)*, **3**(9), 1511–1523.
- Lee, S. and van der Werf, J. (2016). MTG2: An efficient algorithm for multivariate linear mixed model analysis based on genomic information. *Bioinformatics*, **32**, 1420–1422.
- Li, Z. and Sillanpää, M. J. (2013). A Bayesian nonparametric approach for mapping dynamic quantitative traits. *Genetics*, **194**(4), 997–1016.
- Li, Z. and Sillanpää, M. J. (2015). Dynamic quantitative trait locus analysis of plant phenomic data. *Trends in Plant Science*, **20**(12), 822–833.
- Li, Z. *et al.* (2014). Functional multi-locus QTL mapping of temporal trends in Scots pine wood traits. *G3 (Bethesda)*, **4**(12), 2365–2379.
- Misztal, I. *et al.* (2002). BLUPF90 and related programs (BGF90). *Proceedings of the 7th world congress on genetics applied to livestock production*, **33**, 743–744.
- Monterrubio-Gómez, K. *et al.* (2019). Posterior inference for sparse hierarchical non-stationary models. *arXiv e-prints*, page arXiv:1804.01431.
- Moore, C. R. *et al.* (2013). High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics*, **195**(3), 1077–1086.
- Moore, R. *et al.* (2019). A linear mixed-model approach to study multivariate gene–environment interactions. *Nature Genetics*, **51**, 180–186.
- Murray, I. and Adams, R. P. (2010). Slice sampling covariance hyperparameters of latent Gaussian models. *Advances in Neural Information Processing Systems*, **2**, 1732–1740.
- Murray, I. *et al.* (2010). Elliptical slice sampling. *Proceedings of Machine Learning Research*, **9**, 541–548.
- Ni, G. *et al.* (2018). Genotype-covariate correlation and interaction disentangled by a whole-genome multivariate reaction norm model. *Nature Communications*, **10**, 2239–2253.
- Oliveira, H. *et al.* (2019). Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. *Journal of Dairy Science*, **102**(3), 2365–2377.
- Pletcher, S. D. and Geyer, C. J. (1999). The genetic analysis of age-dependent traits: modeling the character process. *Genetics*, **153**, 825–835.
- Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press, Cambridge.
- Robert, C. P. and Casella, G. (2009). *Introducing Monte Carlo Methods with R*. Springer, New York.
- Roberts, G. O. and Rosenthal, J. S. (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, **18**(2), 349–367.
- Robinson, M. *et al.* (2017). Genotype–covariate interaction effects and the heritability of adult body mass index. *Nature Genetics*, **49**, 1174–1181.
- Roininen, L. *et al.* (2014). Whittle–Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography. *Inverse Problems and Imaging*, **8**, 561–586.
- Schaeffer, L. R. (2016). *Random Regression Models*.
- Sillanpää, M. J. (2011). On statistical methods for estimating heritability in wild populations. *Molecular Ecology*, **20**, 1324–1332.
- Sillanpää, M. J. *et al.* (2012). Simultaneous estimation of multiple quantitative trait loci and growth curve parameters through hierarchical Bayesian modeling. *Heredity*, **108**, 134–146.
- Stinchcombe, J. R. *et al.* (2012). Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. *Trends in Ecology & Evolution*, **27**(11), 637–647.
- Sung, Y. *et al.* (2016). A genome-wide association study of multiple longitudinal traits with related subjects. *Stat*, **5**(1), 22–44.
- Vanhatalo, J. *et al.* (2019). A Gaussian process model and Bayesian variable selection for mapping function-valued quantitative traits with incomplete phenotypic data. *Bioinformatics*, **35**(19), 3684–3692.
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, **91**(11), 4414–4423.
- Visscher, P. M. *et al.* (2008). Heritability in the genomics era – concepts and misconceptions. *Nature Reviews Genetics*, **9**, 255–266.
- Xiong, H. *et al.* (2011). A flexible estimating equations approach for mapping function-valued traits. *Genetics*, **189**(1), 305–316.
- Young, A. (2019). Solving the missing heritability problem. *PLOS Genetics*, **15**, e1008222.
- Yu, Y. and Meng, X.-L. (2011). To center or not to center: that is not the question – an ancillarity–sufficiency interweaving strategy (ASIS) for boosting MCMC efficiency. *Journal of Computational and Graphical Statistics*, **20**(3), 531–570.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, **99**(465), 250–261.