Conditional repeatability and the variance explained by reaction norm variation in random slope models

Holger Schielzeth[1], Shinichi Nakagawa[2]

[1]Institute of Ecology and Evolution, Friedrich Schiller University Jena, Germany
[2]Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia

Data and Code availability: There is no data to be deposited. R scripts for simulations are available on Github (https://github.com/hschielzeth/RandomSlopeR2).

Address for correspondence: Holger Schielzeth, Institute of Ecology and Evolution, Friedrich Schiller University Jena, Dornburger Str. 159, 07743 Jena, Germany, holger.schielzeth@uni-jena.de

Abstract

Individuals differ in average phenotypes and in sensitivity to environmental variation. Such context-sensitivity can be modelled as random-slope variation. Random-slope variation implies that the proportion of between-individual variation varies across the range of a covariate (environment/context/time/age) and has thus been called 'conditional' repeatability. We propose to put conditional repeatabilities in perspective of the total phenotypic variance and suggest a way of standardization using the random-slope coefficient of determination R_S^2. Furthermore, we illustrate that the marginalized repeatability R_{mar} averaged across an environmental gradient offers a biologically relevant description of between-individual variation. We provide simple equations for calculating key descriptors of conditional repeatabilities, clarify the difference between random-intercept variation and average between-individual variation and make recommendations for comprehensive reporting. While we introduce the concept with individual variation in mind, the framework is equally applicable to other type of between-group/cluster variation that varies across some (environmental) gradient.

Introduction

Repeatabilities *R* and coefficients of variation R^2 allow a decomposition of the sources of biological variance in some response, feature or trait of interest. Repeatabilities (also known as intra-class correlations, ICC) are concerned with a decomposition of random-effect variances (Nakagawa and Schielzeth 2010, Wolak *et al.* 2012) and have become particularly relevant in the study of labile and repeatedly expressed phenotypes (Bell *et al.* 2009). Coefficients of determination R^2 quantify the variance explained in the fixed part of the model (marginal R^2; *sensu* Nakagawa and Schielzeth 2013). Repeatabilities and coefficients of variation are thus complementary quantities, one focusing on the random the other on the fixed part of the model.

Both, repeatabilities and coefficients of determination, quantify sources of variation in relation to the total variance in a response. To make this more concrete, we want to focus on variance decomposition in a context of the study of phenotypic variation, although the concepts are easily transferred to other systems. Imagine some flexible phenotype of some organism (this may be some physiological, endocrinological or behavioral trait, see e.g. Nespolo and Franco 2007, Bell *et al.* 2009) that has been measured across multiple individuals with repeated observation per individual. Observed phenotypes y_{ij} are thus clustered within individuals *i* with repeated observations *j* per individual. The phenotypic equation represents a variance decomposition model that consists of a mechanistic (fixed effects) and idiosyncratic (random effects) part (Allegue *et al.* 2017):

$$y_{ij} = \alpha + \beta_1 x_{1,ij} + \beta_2 x_{2,ij} + \cdots + u_i + e_{ij}$$

Where i indexes individuals, j indexes observations, the terms βx represent fixed effect predictors x and their slopes β with numbers indexing different predictors (of which there may be more than the two shown here). The effect of the βx terms may be summarized as the linear predictor $\eta_{ij} = \sum \beta_k x_k$ and represents the fixed part of the model. The terms u and e are the random components, where u_i represents deviations of individuals from the population mean and e_{ij} represents deviations of observations from individual means. Individual-level deviations u_i and observation-level deviations e_{ij} are typically assumed to be normally distributed with mean of zero and variance estimated from the data. Since the linear predictor also explains some phenotypic variance, there are three variance components, σ_η^2, σ_u^2 and σ_e^2, that can be interpreted as sources of biological variation.

Various options for estimating ratios of these variance components for the random and the fixed part of the model have been discussed extensively (O'Dea *et al.* 2021). Here we are concerned with another element, random-slope variation that blurs the distinction between the random and the fixed part of the model. Random slopes are interactions between fixed-effect covariates and random-effects: Slopes that vary by random-effect level (Gelman and Hill 2007, Dingemanse and Dochtermann 2013). Random-slopes are often important, because they allow to control for pseudoreplication in the estimation of the population slope (Schielzeth and Forstmeier 2009, Gurka *et al.* 2011). Furthermore, they represent phenotypically plastic responses to an environment and are therefore relevant in the study of reaction norms (Nussey *et al.* 2007, Dingemanse *et al.* 2010). Random-slope models have

become popular in the study of ecology and evolution, because they reflect phenotypic plasticity such as an organisms' ability to cope with environmental changes (Fox *et al.* 2019).

Random-slope variation disrupts the concept of "repeatability" (Biro and Stamps 2015). With random-slopes, the amount of between-individual variation, for example, depends on levels of the covariate, thus on context, environment, age or time. We have therefore introduced the term "conditional repeatability" for repeatabilities that vary by covariates (Nakagawa and Schielzeth 2010). Although random-slope models have become popular, we are not aware of any universal standardized measure of random-slope variation that is in frequent use. A general framework for variance decomposition in mixed-models has recently been proposed in psychology (Rights and Sterba 2019, 2020). We here focus on particular elements of this framework that we think are most applicable in ecology and evolution. Since meta-analyses on the magnitude of random-slope variation are missing, we also know very little about the magnitude of random-slope variation in natural systems. It therefore needs a system for estimating random-slope variation in a manner that is comparable across study system: It needs a method of standardization.

Johnson (2014) has introduced equations to estimate repeatabilities form random slope models. Johnson's method is based on the multiplication of the random-intercept random-slope variance-covariance matrix $\mathbf{\Sigma}$ with the design matrix $\mathbf{X}$ for the fixed effects (called $\mathbf{Z}$ in Johnson 2014). Effectively, Johnson's repeatability estimates average repeatabilities across the range of the covariate. As such, Johnston's repeatability is different from and typically larger than then random-intercept variation estimated in the model. In other words, with random-slope variation, the random-intercept variation is no longer a comprehensive parameter that describes the magnitude of individual differences (or differences among other types of groups). Random-intercept variation merely describes individual-differences at a single point of the environment.

Rights and Sterba (2019, 2020) have introduced a comprehensive system for calculating various R^2 measures from mixed-effects models. Unlike Johnson (2014), they use the variances and covariances of fixed effects rather than the design matrices for the estimation. This offers a more concise version for reporting and it is therefore the approach that we adopt below. However, although Rights and Sterba (2019, 2020) strongly argue for single-source R^2 as the main focus of reporting and interpretation, we think that this does not capture the most biologically relevant estimates. An inherent feature of reaction norms that vary across contexts is that the between-individual variance is variable across contexts and an additive decomposition (as shown in Rights and Sterba 2019, 2020) is no longer most efficient.

We here discuss the concept of conditional repeatabilities in some details. We start with theoretical considerations and then proceed to application. In particular, we present a calculation for variance-standardized random-slope variances and make recommendations for comprehensive reporting. Some of the quantities have been proposed by Rights and Sterba (2019), but their value for variance decomposition in ecology and evolution has not yet been appreciated. One important message of our paper is that the variance in intercepts as estimated in random-slope models is a rather arbitrary value (the random effect variance

at the point where the covariate is zero), but that a biologically more relevant quantity, the average between-group variance across a meaningful range of the covariate can be easily calculated.

Theory

Single random slope with mean-centered covariates

We first consider a phenotypic equation expressed as a mixed effect regression model and assume that all parameters are known with certainty.

$$y_{ij} = \alpha + (\beta + v_i)x_{ij} + u_i + \varphi + e_{ij}$$

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}\right), \boldsymbol{\Sigma} = \begin{bmatrix} V_u & C_{u,v} \\ C_{u,v} & V_v \end{bmatrix}$$

$$e_{ij} \sim N\,(0, V_R)$$

$$x_{ij} \sim D\,(\mu, V_x)$$

Where y_{ij} is the response of interest, α is the global intercept (an estimate of the population mean if covariates are centered), x_{ij} is a covariate (context, environment, age or time) that varies within individuals (an observation-level predictor), β is the population mean response to the covariate, u_i is the deviation of mean individual trait values from the population mean, v_i is the deviation of individual slopes from the population mean slope and e_{ij} are residual deviations. φ is merely introduced as a placeholder for other additive variance components such as additional fixed or random effects. The variance explained by φ is V_φ. The residual deviations e_{ij} are normally distributed with a variance V_R, and u_i as well as v_i are multivariate normal distributed with variances of V_u and V_v, respectively, and a covariance of $C_{u,v}$. The covariate x_{ij} is arbitrarily distributed (as symbolized by *D*) with mean μ and variance of V_x. We will first assume that covariances are grand-mean centered, such that $\mu = 0$ and lift the constraint later.

The equation translates into the following variance components (Snijders and Bosker 2011, Allegue *et al.* 2017, Rights and Sterba 2019):

Variance explained by fixed effects: $V_F = \beta^2 V_x$
Variance explained by individual: $V_I = V_u + V_v V_x$
Other variance components: V_φ
Residual variance: V_R
Phenotypic variance: $V_P = V_F + V_I + V_\varphi + V_R$

V_I refers to the total variance explained by individual identity, including random intercept and random slope variance. This is an interesting quantity that summarizes individual

differences, an important topic of current research (Réale *et al.* 2007, Stamps 2016), but it is different from (generally larger than) V_u if random-slope variation differs from zero. We also call V_I the marginalized variance explained by individual, because it averages (marginalizes) across the range of the covariate.

We note that the phenotypic variance V_P as we calculate it here as the sum of additive variance components might differ slightly from the variance in response values as estimated from the raw data (Rights and Sterba 2020). The difference is that the sum of the variance components aims to estimate the population variance while the variance in raw response values represents to variance in the sample. Since the population variance is what is relevant to biological interpretation (de Villemereuil *et al.* 2018), the sum of additive components is usually preferable. However, if components are not fully additive, this may lead to misestimation. It is therefore important to specify the variance decomposition model correctly.

One component of the total between-individual variation V_I is the variance uniquely explained by random slopes. This is the phenotypic variation explained by individual differences in reaction norms:

$$V_S = V_v V_x$$

Even though V_S is contributing to V_I, there is no decomposition of V_I in additive components (see below).

Single random slope with uncentered covariates

We have so far assumed that $\mu = 0$, which can be easily achieved by mean-centering covariates prior to the analysis (note that with z-transformation of x, $V_S = V_v$). Covariate centering is generally advisable when used in random-slope models, because uncentered covariates tend to produce large covariation between random-slopes and random-intercepts, which often leads to convergence problems in model fitting. If the covariate was not centered, then

$$V_S = V_v V_x + \mu^2 V_v$$

$$V_I = V_u + V_v V_x + \mu^2 V_v + 2\mu C_{u,v}$$

The between-individual variation in average phenotypes is a little more difficult, because as a conditional repeatability it varies across the range of the covariate. We can calculate the amount of between-individual variation for any point x as (Figure 1):

$$V_{I,x} = V_u + x^2 V_v + 2x C_{u,v}$$

The minimum value of $V_{I,x}$ (from where it increases in either direction) is reached at:

$x_{min} = \frac{-C_{u,v}}{V_v}$.

It follows that the minimum value of $V_{I,x}$ is:

$$min(V_{I,x}) = V_u - \frac{C_{u,v}^2}{V_v}.$$

With context-sensitive responses, it is difficult to conceptualize a pure among-individual variation in elevation across the entire gradient (that is, there is no additive decomposition of V_I). We think that V_I is still the best descriptor of overall magnitude in individual differences (Figure 1). One might be tempted to use $V_{I*} = V_I - V_S = V_u$ as an estimator of elevation, but this is just the between-individual variance at the point where the covariate is zero. Whether this is a meaningful value, depends on how the covariate is centered. The value might be representative for an average covariate value with mean-centered covariates. However, whether or not this is also the minimum value of between-individual variation depends on the intercept-slope covariance that, as a property of the population, is usually beyond experimental control.

Multiple and correlated predictors

We have above shown equations that use means and variances of covariate x to quantify conditional repeatabilities. Johnson (2014) had introduced a different approach that uses the specific design matrix $\mathbf{X}$ for the fixed effects instead. This approach will be particularly useful when fixed effect predictors are correlated, since positive correlations will inflate the contribution of a predictor to the phenotypic variance beyond $\beta^2 V_x$.

$$V_I = Tr(\mathbf{X}\Sigma\mathbf{X}')/n$$

Where $\mathbf{X}$ is the model matrix for the intercept and the fixed effect of interest (typically a column of 1s for the intercept and columns of covariate values x for all observations), $\mathbf{X}'$ is the transpose of $\mathbf{X}$, n is the total number of observations and $\mathbf{\Sigma}$ is the random intercept-slope covariance matrix as defined above and as estimated from the data, Tr signifies the trace (the sum of the diagonal elements) of the resultant square matrix. To put it simply, V_I is calculated as the predicted amount of between-individual variation associated with all observations and averaged across the dataset.

Johnson's (2014) approach can be useful for computation, but the alternative use of means and variances of covariates makes reporting much easier. Our simulations (see below) show that means and variances are no less accurate, and yield unbiased estimates in most cases. Furthermore, it is also possible to estimate the variance in the fixed part explained by correlated predictors as:

$$V_F = \beta_1^2 V_{x_1} + \beta_2^2 V_{x_2} + 2\beta_1\beta_2 C_{x_1,x_2}$$

$$\boldsymbol{\Omega} = \begin{bmatrix} V_{x_1} & C_{x_1,x_2} \\ C_{x_1,x_2} & V_{x_2} \end{bmatrix}$$

for a model with two fixed effect predictors x_1 and x_2 with variances of V_{x_1} and V_{x_2}, respectively and predictor covariance of C_{x_1,x_2}. The terms β_1 and β_2 represent the respective population-level slopes. The correlation among predictors is thus summarized in the predictor variance-covariance matrix $\boldsymbol{\Omega}$. The predictor variance-covariance matrix $\boldsymbol{\Omega}$ is suitable for concise reporting and we suggest that it can be reported as part of the supplementary materials with experimental studies.

Multiple random slopes with uncentered covariates

We have so far assumed that φ is an arbitrary, but additive effect. This might include other random-intercept components or additional fixed effects, which are easy to deal with. In general, fixed and random-effects might vary between individuals, such as any group-level fixed effect (like the size of an individual) or higher-level random effect (like the family or patch that an individual originates from). The equations for calculating V_I thus apply to the random-effect variances associated with an individual identity random effect only, that is, after controlling for other between-individual effects in the model. This is generally the case for repeatabilities and might thus be quite familiar.

A different complication arises when additional fixed effects are also involved in random slopes that vary between individuals. This makes the calculations more tedious, but it is still possible to calculate the total variance explained by individual identity across multiple covariates. For example, if there are two covariates x_1 and x_2 that are involved in an interaction with the individual random-effect component to yield random-slopes $v_{1,i}$ and $v_{2,i}$.

$$y_{ij} = \alpha + (\beta_1 + v_{1,i})x_{1,ij} + (\beta_2 + v_{2,i})x_{2,ij} + u_i + \varphi + e_{ij}$$

If the means of the two covariates x_1 and x_2 are μ_1 and μ_2, respectively, and the variances in covariates are V_{x_1} and V_{x_2}, respectively, and the covariance between x_1 and x_2 is $C_{x_1x_2}$, than the full variance-covariance structure is $\boldsymbol{\Omega}$ as above. Note that one of the covariates might be the square of the other $x_2 = x_1^2$, such that all calculations can be used for polynomial model fits, too.

The covariance structure for the random intercepts and slopes is:

$$\begin{bmatrix} u_i \\ v_{1,i} \\ v_{2,i} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}\right), \boldsymbol{\Sigma} = \begin{bmatrix} V_u & C_{u,v_1} & C_{u,v_2} \\ C_{u,v_1} & V_{v_1} & C_{v_1,v_2} \\ C_{uv_2} & C_{v_1v_2} & V_{v_2} \end{bmatrix}$$

Where V_u represents the random-intercept variance, V_{v_1} and V_{v_2} are the random-slope variances for covariates x_1 and x_2 and covariances are shown in the off-diagonal. The total variance explained by the individual random effect can then be calculated as:

$$V_I = V_u + V_{v_1}V_{x_1} + V_{v_2}V_{x_2} + \mu_1^2 V_{v_1} + \mu_2^2 V_{v_2} + 2\mu_1 C_{u,v_1} + 2\mu_2 C_{u,v_2} + 2C_{v_1x_1,v_2x_2}$$

The covariance between the products of random numbers can be calculated as $cov(AC, BD) = \mathrm{E}(A)\mathrm{E}(B)cov(C, D) + \mathrm{E}(C)\mathrm{E}(D)cov(A, B) + cov(A, B)cov(C, D)$. In the case of random slopes that are estimated as deviations form the mean, the expectation for $v_{1,i}$ and $v_{2,i}$ is zero $[\mathrm{E}(v_1) = \mathrm{E}(v_2) = 0]$, such that $C_{v_1x_1,v_2x_2} = \mu_1\mu_2 C_{v_1,v_2} + C_{v_1v_2}C_{x_1,x_2}$. Therefore, the full equation for variance explained by the individual random effect is:

$$V_I = V_u + V_{v_1}V_{x_1} + V_{v_2}V_{x_2} + \mu_1^2 V_{v_1} + \mu_2^2 V_{v_2} + 2\mu_1 C_{u,v_1} + 2\mu_2 C_{u,v_2} \\ +2(\mu_1\mu_2 C_{v_1,v_2} + C_{v_1,v_2}C_{x_1,x_2})$$

Which can be rearranged to:

$$V_I = V_u + V_{v_1}(V_{x_1} + \mu_1^2) + V_{v_2}(V_{x_2} + \mu_2^2) + 2(\mu_1 C_{u,v_1} + \mu_2 C_{u,v_2}) + 2C_{v_1,v_2}(\mu_1\mu_2 + C_{x_1,x_2})$$

While these calculations seem long and tedious, they can be easily implemented in scripted analysis even for complex models.

Standardization of variances

The above equations offer obvious ways of variance-standardization for both total between-individual variation V_I and random-slope variation V_S. We propose to call the variance-standardized average between-individual variation V_I the marginalized repeatability R_{mar}, because it marginalizes (averages) across the environmental gradient:

$$R_{mar} = \frac{V_I}{V_P} = \frac{V_u + V_v V_x + \mu^2 V_v + 2\mu C_{u,v}}{V_F + V_I + V_\varphi + V_R}$$

This value is typically larger than the variance-standardized random-intercept variation. We propose to call the variance-standardized random-slope variation the random-slope coefficient of determination R_S^2 (called $R_t^{2(v)}$ in Rights and Sterba 2019).

$$R_S^2 = \frac{V_S}{V_P} = \frac{V_v V_x + \mu^2 V_v}{V_F + V_I + V_\varphi + V_R}$$

Variance-standardization puts the variance explained by individual components in perspective of the total phenotypic variance, which, in our experience, is what ecologists and evolutionary biologists are usually interested in (see de Villemereuil *et al.* 2018 for a discussion). However, it has been argued that variance-standardization may produce different values not because of differences in the numerator, but because of differences in the denominator (Houle 1992). An alternative way of standardization is therefore standardization by the square of the mean trait value, if the trait is ratio-scale and the assumption that the variance explained increases with the square of the mean is reasonable

(Houle *et al.* 2011). For mean-standardization, V_P has to be replaced by $\bar{y}^2$ (or α^2 if all covariates are mean-centered) in the equations above.

Reporting

We write this article partly to encourage complete reporting for future meta-analyses of phenotypic plasticity. A first important message is that the mean of and variance in the covariate are important quantities that allow putting reaction norms and reaction norm variation in perspective of the phenotypic variance. One way to standardize random slopes is the use of variance-standardized covariates $x' = \frac{x-\bar{x}}{\sigma_x}$, in which case $\bar{x}' = 0$ and $V_{X'} = 1$. Alternatively, or better additionally, raw mean $\bar{x}$ and variance V_x should be reported. Furthermore, it is important to report covariation among fixed effect, ideally in the form of a predictor variance-covariance matrix $\mathbf{\Omega}$. A second important message is that the correlation among random slopes and random intercepts is an important parameter that can be biologically interpreted and should be reported. Negative correlations show that between-individual variation is lower at high values of the covariate and *vice versa* for positive correlations.

For meta-analysis, it would also need some measures of uncertainty in all relevant estimates. This could be easily achieved by applying our equations to samples from the posterior distributions of models fit in a Bayesian framework (Gelman *et al.* 2004). In a likelihood framework, it could be achieved by parametric bootstrapping (Faraway 2014). However, some of the sampling variances are small in comparison to others and might not need to be available with full uncertainty estimates. For example, estimates of covariate means and variances are estimated with far higher precision than estimates of random effect variances. In fact, it might sometimes be useful to use means and variances for environmental covariates from independent data if the data were collected in an experimental setting where variance in the covariate was manipulated.

Simulations

We have implemented simple simulation to illustrate some critical point. We do not aim to present a full exploration of the full parameter space, that is potentially vast and varies between applications. For simulations that explore the power of different sampling designs for estimating for estimating random-slope variation we refer to Martin *et al.* (2011), van de Pol (2012) or Westneat *et al.* (2020).

We implemented a data-generating function for a simple random-slope model with a single grouping factor and two covariates. Random-slopes act on an observation-level predictor x. Furthermore, a group-level covariate φ was introduced with an associated slope γ. Random-slopes and random-intercepts were generated from a multivariate normal distribution with means of zero and variance-covariance matrix $\mathbf{\Sigma}$. The generating phenotypic equation as used in our simulations was:

$$y_{ij} = \alpha + (\beta + v_i)x_{ij} + u_i + \gamma \cdot \varphi_{ij} + e_{ij}$$

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}\right), \boldsymbol{\Sigma} = \begin{bmatrix} V_u & C_{u,v} \\ C_{u,v} & V_v \end{bmatrix}$$

$$e_{ij} \sim N\,(0, V_R)$$

$$x_{ij} \sim N\,(\mu_x, V_x)$$

$$\varphi_{ij} \sim N\,(\mu_\varphi, V_\varphi)$$

We used the following parameter settings: $\alpha = 3$, $\beta = -0.5$, $\gamma = 0.5$, $V_R = 1$, $V_u = 1$, $V_v = 0.5$, $C_{u,v} = 0.3$, $V_x = 1.2$, $V_\varphi = 1$, $\mu_x = 0.5$, $\mu_\varphi = -0.5$. These values were relatively arbitrarily chosen, and effects are purposefully rather strong in order to demonstrate general patterns. The broad patterns are largely insensitive to the detailed choice of values. Covariate values x were drawn from uniform distributions. We simulated a population of 60 individuals with an average of 10 observations per individual, which we consider a moderate sample size. For each of 500 simulation runs we fitted the regression model, estimated the parameters and used the above equations to quantify important parts of conditional repeatability. Data were generated in R 4.1.2 (R Core Team 2021) and analyzed using random-slope models fitted in lme4 1.1-27.1 (Bates *et al.* 2015).

The basic setting showed that the conditional repeatability was quite accurately estimated, with only minor bias in x_{min} and marginal downward bias in V_I (Figure 2a). The accuracy in the estimation of model parameters and derived quantities is unequally distributed with some (like α, μ_x, V_x and V_R) being estimated with high accuracy while others (like $C_{u,v}$, V_F, V_φ and x_{min}) being estimated with much less accurately (Figure 3). We also simulated a reduced sample size of 30 individuals and an average of 3 observations per individual. The small-sample simulation resulted substantially increased variation in estimates, but no strong bias except for pronounced bias in x_{min} (Figure 2b).

We then assessed estimates when only random intercepts and no random slopes were fitted to the same data (note that data were generated with random-slope variation). These models do not estimate conditional repeatabilities, but rather an overall repeatability. The random-intercept variance was substantially larger than the V_u that was simulated, hence some reaction norm variation in the data was converted to between-individual variation in intercepts in the model. Yet while the estimated was intermediate between V_u and V_I, it substantially underestimated the total individual variance V_I as determined by the phenotypic equation (Figure 4a). Hence, if there is significant reaction norm variation in the population, these models thus typically misestimate average between-individual variation as well as individual variation at the point where the covariate is zero.

We further explored the effect of measurement error in the covariate on estimates of individual components. Linear models assume that covariates are measured without error (Snijders and Bosker 2011), hence this might sound like a non-sensible attempt. However,

measurement error is inevitable, for example when within-subject centering is used to separate within and between-individual responses to some covariate (Raudenbush and Bryk 2002, van de Pol and Wright 2009). Simulations with the same moderate sample size as above and a rather large measurement error of 30% in the covariate resulted in an underestimation of reaction norm variation V_S and underestimation of total between-individual variation V_I, but an overestimation of V_u (Figure 4b). Hence, some of the reaction norm variation was converted to random-intercept variation. This is sometimes unavoidable as in the case of within-subject centering. If the amount of error can be estimated, it could, in principle, be corrected for.

Conclusions

We present equations that allow the description of conditional between-individual variances (based on Johnson 2014, Rights and Sterba 2019, 2020). Most importantly, we introduce a way of standardized reporting of reaction norm variation, clarify the difference between random-intercept variation and average (marginalized) between individual variation and make recommendations for comprehensive reporting. By putting reaction norm variation in perspective of the phenotypic variance, we aim to promote comprehensive variance decomposition of natural variation in traits of interest. We hope that these tools will stimulate more research on context-sensitivity of individual (or other group-level) variation and will allow data for future meta-analyses to accumulate.

Acknowledgements

HS were supported by the German Research Foundation (DFG) as part of the SFB TRR 212 (NC³) (funding INST 215/543-1, 396782608). SN was supported by the ARC Discovery Project grant (DP180100818).

Data availability statement

There is no data to be deposited. R scripts for simulations are available on https://github.com/hschielzeth/RandomSlopeR2. The site also holds two R function condR and condRplot that can be used for custom calculations and plotting of conditional variances.

Author contributions

HS conceived the idea, implemented the simulations and drafted the manuscript; SN helped to refine the method, contributed mathematical derivations and revised the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

References

Aikens, E. O., M. J. Kauffman, J. A. Merkle, S. P. H. Dwinnell, G. L. Fralick, and K. L. Monteith (2017) The greenscape shapes surfing of resource waves in a large migratory herbivore. *Ecology Letters*, 20, 741-750.

Allegue, H., Y. G. Araya-Ajoy, N. J. Dingemanse, N. A. Dochtermann, L. Z. Garamszegi, S. Nakagawa, D. Réale, H. Schielzeth, and D. F. Westneat (2017) Statistical Quantification of Individual Differences (SQuID): an educational and statistical tool for understanding multilevel phenotypic data in linear mixed models. *Methods in Ecology and Evolution*, 8, 257-267.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1-48.

Bell, A. M., S. J. Hankison, and K. L. Laskowski (2009) The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, 77, 771-783.

Biro, P. A. and J. A. Stamps (2015) Using repeatability to study physiological and behavioural traits: ignore time-related change at your peril. *Animal Behaviour*, 105, 223-230.

Bolund, E., H. Schielzeth, and W. Forstmeier (2009) Compensatory investment in zebra finches: Females lay larger eggs when paired to sexually unattractive males. *Proceedings of the Royal Society B*, 276, 707–715.

Clough, Y., J. Ekroos, A. Baldi, P. Batary, R. Bommarco, N. Gross, A. Holzschuh, S. Hopfenmuller, E. Knop, M. Kuussaari, R. Lindborg, L. Marini, E. Ockinger, S. G. Potts, J. Poyry, S. P. M. Roberts, I. Steffan-Dewenter, and H. G. Smith (2014) Density of insect-pollinated grassland plants decreases with increasing surrounding land-use intensity. *Ecology Letters*, 17, 1168-1177.

Craven, D., N. Eisenhauer, W. D. Pearse, Y. Hautier, F. Isbell, C. Roscher, M. Bahn, C. Beierkuhnlein, G. Bönisch, N. Buchmann, C. Byun, J. A. Catford, B. E. L. Cerabolini, J. H. C. Cornelissen, J. M. Craine, E. De Luca, A. Ebeling, J. N. Griffin, A. Hector, J. Hines, A. Jentsch, J. Kattge, J. Kreyling, V. Lanta, N. Lemoine, S. T. Meyer, V. Minden, V. Onipchenko, H. W. Polley, P. B. Reich, J. van Ruijven, B. Schamp, M. D. Smith, N. A. Soudzilovskaia, D. Tilman, A. Weigelt, B. Wilsey, and P. Manning (2018) Multiple facets of biodiversity drive the diversity–stability relationship. *Nature Ecology & Evolution*, 2, 1579-1587.

de Villemereuil, P., M. Morrissey, S. Nakagawa, and H. Schielzeth (2018) Fixed effect variance and the estimation of repeatabilities and heritabilities: Issues and solutions. *Journal of Evolutionary Biology*, 31, 621-632.

Dingemanse, N. J. and N. A. Dochtermann (2013) Quantifying individual variation in behaviour: mixed-effect modelling approaches. *Journal of Animal Ecology*, 82, 39-54.

Dingemanse, N. J., A. J. Kazem, D. Réale, and J. Wright (2010) Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology & Evolution*, 25, 81-89.

Faraway, J. J. (2014) *Linear models in R*. 2nd edn. CRC Press, Boca Raton.

Fox, R. J., J. M. Donelson, C. Schunter, T. Ravasi, and J. D. Gaitan-Espitia (2019) Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Philosophical Transactions of the Royal Society B*, 374, 20180174.

Fricke, E. C. and S. J. Wright (2016) The mechanical defence advantage of small seeds. *Ecology Letters*, 19, 987-991.

Gelman, A., J. B. Carlin, H. S. Stern, and D. R. Rubin (2004) *Bayesian Data Analysis*. Chapman & Hall, Boca Raton.

Gelman, A. and J. Hill (2007) *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge, U.K.

Grossenbacher, D., R. B. Runquist, E. E. Goldberg, and Y. Brandvain (2015) Geographic range size is predicted by plant mating system. *Ecology Letters*, 18, 706-713.

Gurka, M. J., L. J. Edwards, and K. E. Muller (2011) Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine*, 30, 2696-2707.

Houle, D. (1992) Comparing evolvability and variability of quantitative traits. *Genetics*, 130, 195-204.

Houle, D., C. Pelabon, G. P. Wagner, and T. F. Hansen (2011) Measurement and meaning in biology. *Quarterly Review of Biology*, 86, 3-34.

Johnson, P. C. D. (2014) Extension of Nakagawa & Schielzeth's R^{2GLMM} to random slopes models. *Methods in Ecology and Evolution*, 5, 944-946.

Krauss, J., R. Bommarco, M. Guardiola, R. K. Heikkinen, A. Helm, M. Kuussaari, R. Lindborg, E. Ockinger, M. Partel, J. Pino, J. Poyry, K. M. Raatikainen, A. Sang, C. Stefanescu, T. Teder, M. Zobel, and I. Steffan-Dewenter (2010) Habitat fragmentation causes immediate and time-delayed biodiversity loss at different trophic levels. *Ecology Letters*, 13, 597-605.

Martin, J. G. A., D. H. Nussey, A. J. Wilson, and D. Réale (2011) Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. *Methods in Ecology and Evolution*, 2, 362-374.

Martin, T. E., R. Ton, and A. Niklison (2013) Intrinsic vs. extrinsic influences on life history expression: metabolism and parentally induced temperature influences on embryo development rate. *Ecology Letters*, 16, 738-745.

Mathot, K. J., P. J. van den Hout, T. Piersma, B. Kempenaers, D. Reale, and N. J. Dingemanse (2011) Disentangling the roles of frequency-vs. state-dependence in generating individual differences in behavioural plasticity. *Ecology Letters*, 14, 1254-1262.

Nakagawa, S. and H. Schielzeth (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, 85, 935-956.

Nakagawa, S. and H. Schielzeth (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133-142.

Nespolo, R. F. and M. Franco (2007) Whole-animal metabolic rate is a repeatable trait: a meta-analysis. *Journal of Experimental Biology*, 210, 3877-3878.

Nussey, D. H., A. J. Wilson, and J. E. Brommer (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology*, 20, 831-844.

O'Dea, R., D. W. A. Noble, and S. Nakagawa (2021) Unifying individual differences in personality, predictability, and plasticity: a practical guide. *Methods in Ecology and Evolution*, doi: 10.1111/2041-210X.13755.

R Core Team (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Raudenbush, S. W. and A. S. Bryk (2002) *Hierarchical linear models: applications and data analysis methods*. 2nd edn. Sage Publications, London, U.K.

Réale, D., S. M. Reader, D. Sol, P. T. McDougall, and N. J. Dingemanse (2007) Integrating animal temperament within ecology and evolution. *Biological Reviews*, 82, 291-318.

Rights, J. D. and S. K. Sterba (2019) Quantifying explained variance in multilevel models: an integrative framework for defining R-squared measures. *Psychological Methods*, 24, 309-338.

Rights, J. D. and S. K. Sterba (2020) New recommendations on the Use of R-squared differences in multilevel model comparisons. *Multivariate Behavioral Research*, 55, 568-599.

Schielzeth, H. and W. Forstmeier (2009) Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology*, 20, 416-420.

Snijders, T. A. B. and R. J. Bosker (2011) *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. 2nd edn. Sage, London.

Stamps, J. A. (2016) Individual differences in behavioural plasticities. *Biological Reviews*, 91, 534-567.

van de Pol, M. (2012) Quantifying individual variation in reaction norms: how study design affects the accuracy, precision and power of random regression models. *Methods in Ecology and Evolution*, 3, 268-280.

van de Pol, M. V. and J. Wright (2009) A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behaviour*, 77, 753-758.

Westneat, D. F., Y. G. Araya-Ajoy, H. Allegue, B. Class, N. Dingemanse, N. A. Dochtermann, L. Z. Garamszegi, J. G. A. Martin, S. Nakagawa, D. Réale, and H. Schielzeth (2020) Collision

between biological process and statistical analysis revealed by mean centring. *Journal of Animal Ecology*, 89, 2813-2824.
Wolak, M. E., D. J. Fairbairn, and Y. R. Paulsen (2012) Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, 3, 129-137.

Tables

Table 1: Examples of random slope models in ecological studies with potential application of conditional repeatabilities.

Topic	**Application**	**Reference**
Functional biodiversity effects on ecosystem stability in grasslands	Spatio-temporal (between study) variation in the dependency of ecosystem stability on species richness	Craven *et al.* 2018
Community composition and pollination in plants	Between-plot variation in the proportion of insect-pollinated plants as it varies across a gradient and landscape features	Clough *et al.* 2014
Habitat fragmentation and biodiversity loss in plants and butterflies	Spatial (between-country) variation in species richness as it depends on past and present landscape features	Krauss *et al.* 2010
Determinants of range size in plants	Between-clade variation in range size of plants as it varies by mating system	Grossenbacher *et al.* 2015
Functional seed ecology of tropical plants	Between-species and between-sample (spatial) variation in seed protective tissue across a gradient of seed sizes	Fricke and Wright 2016
Individual differences in antipredator behavior of red knots	Between-individual variation in anti-predator behavior as it varies across a gradient of predation hazard	Mathot *et al.* 2011
Parental effects on embryo metabolic rates in bird eggs	Between-species variation in the dependency of egg metabolic activity across a gradient of incubation temperatures	Martin *et al.* 2013
Maternal resource allocation in zebra finches	Between-females variation in the dependency of maternal investment into the eggs across a gradient of partner attractiveness	Bolund *et al.* 2009
State variable and resource use in red deer	Between-year variation in synchrony between resource availability and resource use as it depends on individual experience and condition	Aikens *et al.* 2017

Figure legends

Figure 1. Conceptual display of conditional repeatabilities. The figure shows the between individual variance as it depends on the value of the covariate. The random-intercept variance V_u is always estimated at a covariate value of zero. The minimum between-individual variance $min(V_{I,x})$ is reached at x_{min}. The average between-individual variance is V_I and is usually larger than V_u if there is random-slope variation.

Figure 2. Estimation error in conditional repeatabilities with (a) moderate ($N_{ind} = 60$, $N_{obs} = 600$) and (b) small ($N_{ind} = 30$, $N_{obs} = 90$) sample size. The figure shows the same quantities as Figure 1 with the data-generating (true) values shown black. Predicted conditional repeatabilities from 500 replications are shown in grey. Estimated V_I, V_u and $x_{min}/min(V_{I,x})$ are shown in orange, blue and green, respectively, with thin lines representing single iterations and bold lines average values.

Figure 3. Proportional bias in estimates for various model estimates and derived parameters for a base simulation setting (see main text) with 500 replicates. Results are broadly similar for a range of parameter values when the analysis equation matches with the data generation process. The interquartile range is shown in dark shading, the 95% envelope in moderate shading and the total range in light shading. Dots represent individual estimates from 500 replicates and red lines show mean values. Sim = Expectations based on simulation settings, Est = Estimation based on model fit.

Figure 4. Estimation error in conditional repeatabilities (a) when the analysis model is fitted without random slopes and (b) when there is error in the covariate x (30% in this case). The figure shows the same quantities as Figure 1 with the data-generating (true) values shown black. Predicted conditional repeatabilities from 500 replications are shown in grey. Estimated V_I, V_u and $x_{min}/min(V_{I,x})$ are shown in orange, blue and green, respectively, with thin lines representing single iterations and bold lines average values. The case of missing random-slopes does not estimate conditional repeatabilities, but the estimated between-individual variance V_u closely approaches the simulated average between-individual variance V_I.

Figure 1

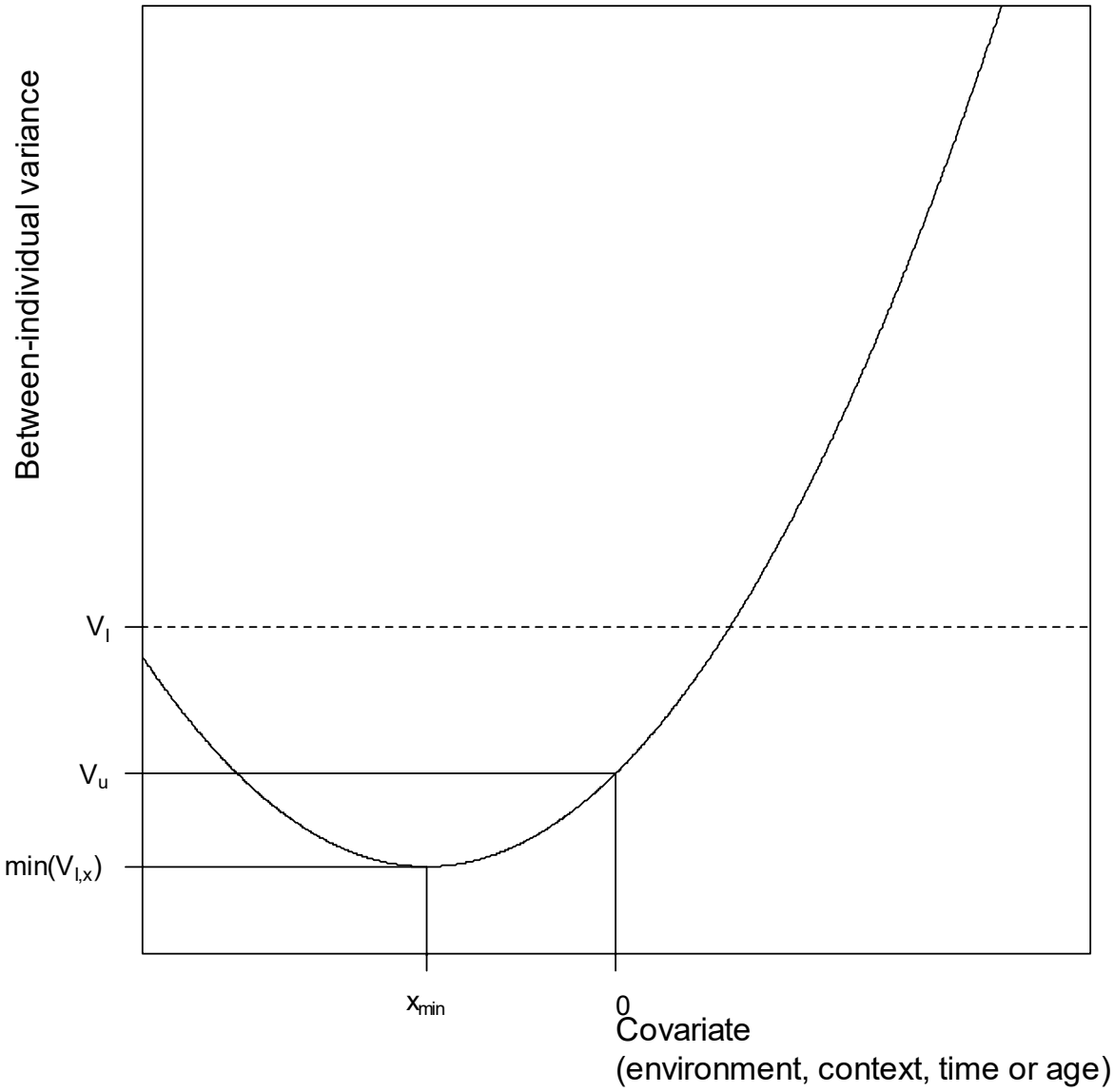
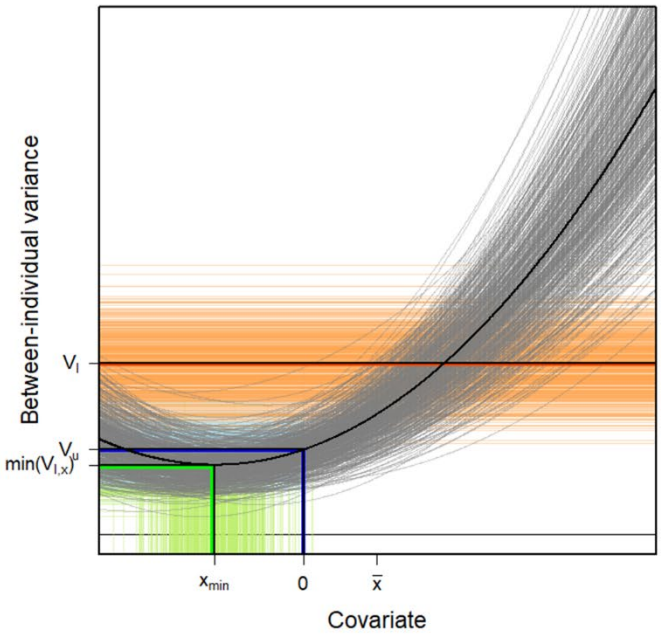

Between-individual variance
V_I
V_u
$min(V_{I,x})$
x_{min}
0
Covariate
(environment, context, time or age)

Figure 2

(a) Moderate sample size



(b) Small sample size

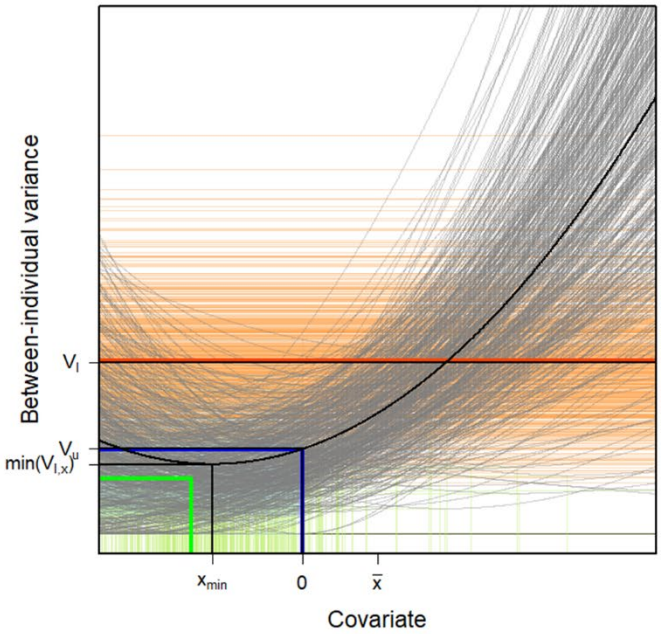


Figure 3

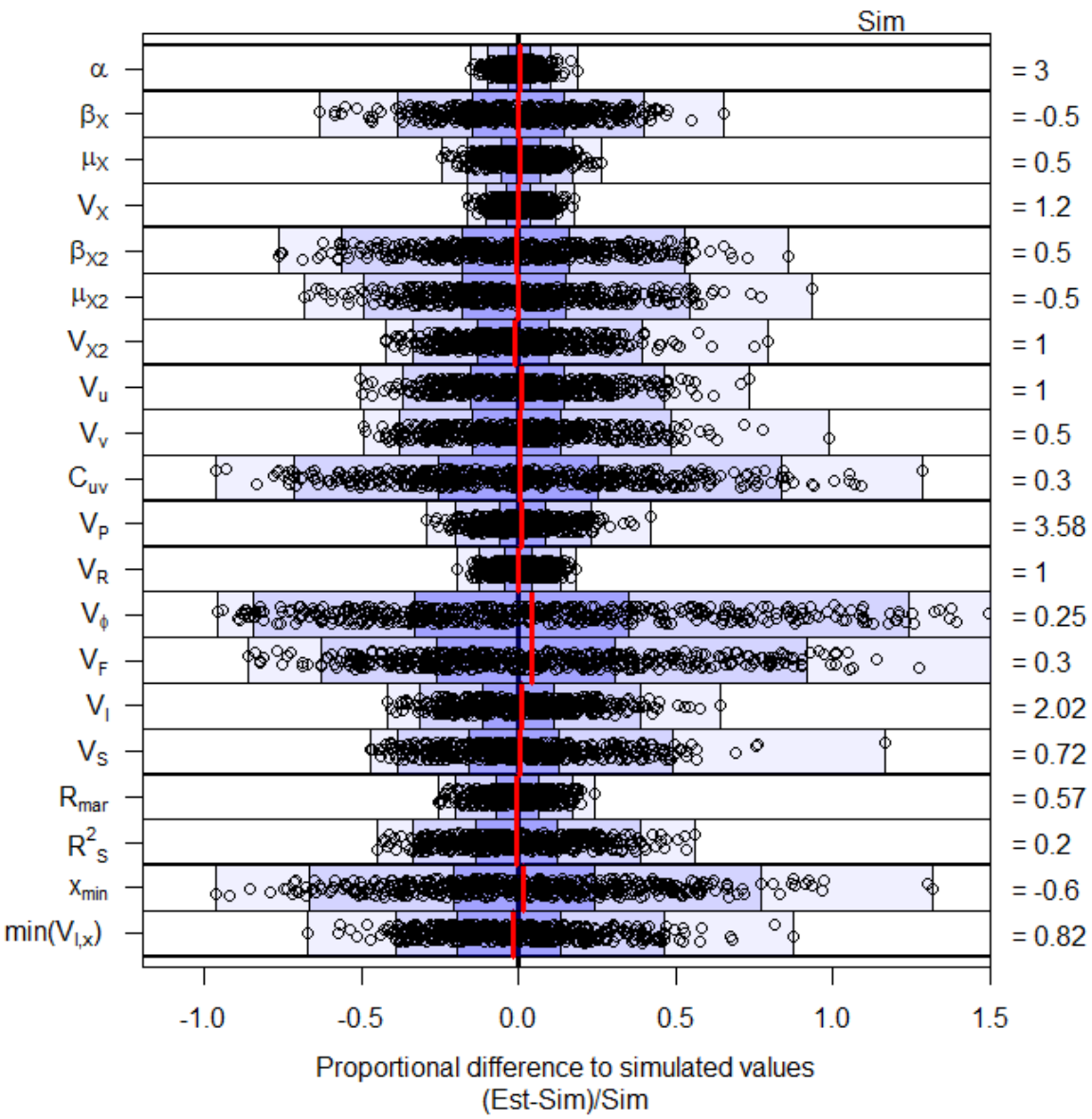
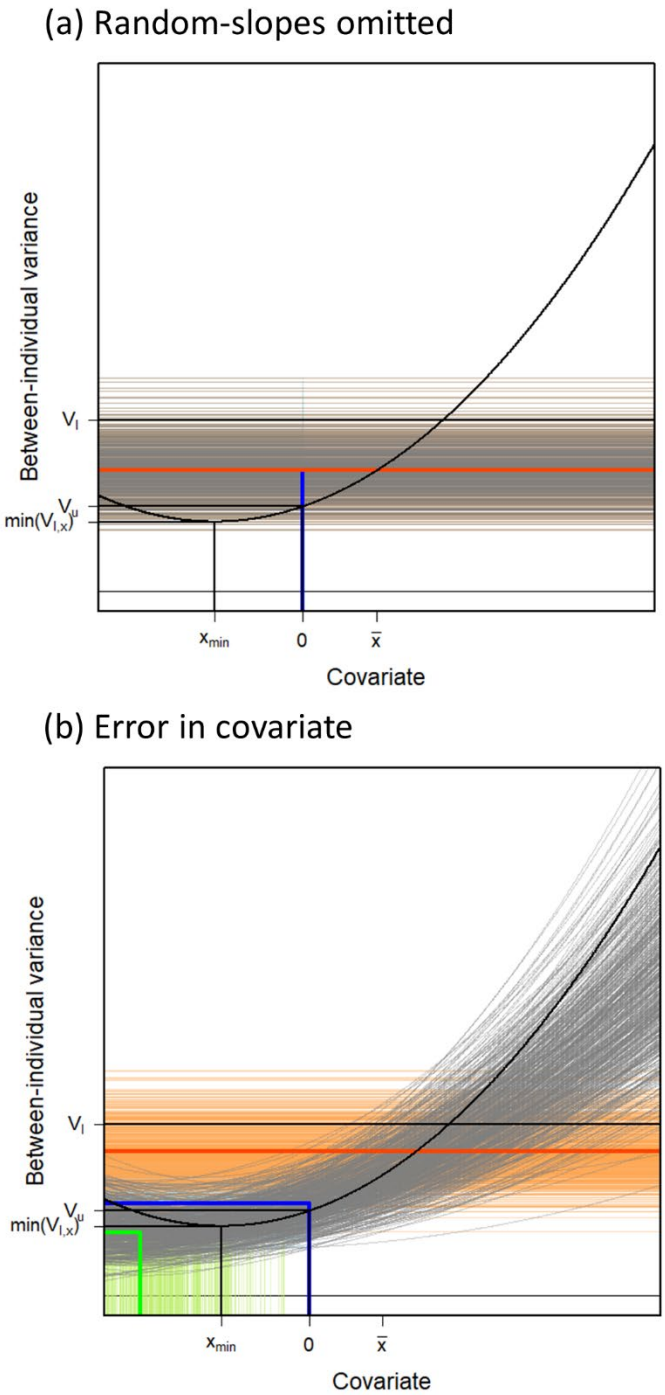

Sim
α
βX
μX
VX
βX2
μX2
VX2
Vu
Vv
Cuv
VP
VR
Vϕ
VF
VI
VS
Rmar
R2S
xmin
min(VI,x)
= 3
= -0.5
= 0.5
= 1.2
= 0.5
= -0.5
= 1
= 1
= 0.5
= 0.3
= 3.58
= 1
= 0.25
= 0.3
= 2.02
= 0.72
= 0.57
= 0.2
= -0.6
= 0.82
-1.0
-0.5
0.0
0.5
1.0
1.5
Proportional difference to simulated values
(Est-Sim)/Sim

Figure 4


(a) Random-slopes omitted
Between-individual variance
V_I
V_μ
$\min(V_{I,x})$
x_{min}
0
$\bar{x}$
Covariate
(b) Error in covariate
Between-individual variance
V_I
V_μ
$\min(V_{I,x})$
x_{min}
0
$\bar{x}$
Covariate