

RESEARCH ARTICLE

Open Access



Multiple-trait structured antedependence model to study the relationship between litter size and birth weight in pigs and rabbits

Ingrid David^{1*} , Hervé Garreau¹, Elodie Balmisse², Yvon Billon³ and Laurianne Canario¹

Abstract

Background: Some genetic studies need to take into account correlations between traits that are repeatedly measured over time. Multiple-trait random regression models are commonly used to analyze repeated traits but suffer from several major drawbacks. In the present study, we developed a multiple-trait extension of the structured antedependence model (SAD) to overcome this issue and validated its usefulness by modeling the association between litter size (LS) and average birth weight (ABW) over parities in pigs and rabbits.

Methods: The single-trait SAD model assumes that a random effect at time t_j can be explained by the previous values of the random effect (i.e. at previous times). The proposed multiple-trait extension of the SAD model consists in adding a cross-antedependence parameter to the single-trait SAD model. This model can be easily fitted using ASReml and the OWN Fortran program that we have developed. In comparison with the random regression model, we used our multiple-trait SAD model to analyze the LS and ABW of 4345 litters from 1817 Large White sows and 8706 litters from 2286 L-1777 does over a maximum of five successive parities.

Results: For both species, the multiple-trait SAD fitted the data better than the random regression model. The difference between AIC of the two models (AIC_random regression - AIC_SAD) were equal to 7 and 227 for pigs and rabbits, respectively. A similar pattern of heritability and correlation estimates was obtained for both species. Heritabilities were lower for LS (ranging from 0.09 to 0.29) than for ABW (ranging from 0.23 to 0.39). The general trend was a decrease of the genetic correlation for a given trait between more distant parities. Estimates of genetic correlations between LS and ABW were negative and ranged from -0.03 to -0.52 across parities. No correlation was observed between the permanent environmental effects, except between the permanent environmental effects of LS and ABW of the same parity, for which the estimate of the correlation was strongly negative (ranging from -0.57 to -0.67).

Conclusions: We demonstrated that application of our multiple-trait SAD model is feasible for studying several traits with repeated measurements and showed that it provided a better fit to the data than the random regression model.

Background

In genetic studies, many traits of interest are repeatedly measured over time, which gives rise to longitudinal data. The main issue when modeling such data is to account for the covariance structure of the repeated records with a

limited number of parameters. A good approach to reduce the number of parameters that need to be estimated (compared with a multiple-trait model) is to use a repeatability model, which is often used because of its simplicity. However, this model assumes, too narrowly, that repeated records are expressions of the same genetic trait and that the phenotypic correlation between repeated measures is uniform. More flexible approaches have been proposed, such as random regression (RR) [1–3], character process (CP), and structured antedependence (SAD) models [4,

*Correspondence: ingrid.david@inra.fr

¹ GenPhySE, INRA, INPT, ENVT, Université de Toulouse, 31326 Castanet-Tolosan, France

Full list of author information is available at the end of the article

5]. The most commonly used approach is the RR model [6], although this approach suffers from various drawbacks, the main one being the well-known “border effect” problem [7]. CP and SAD models have been shown to fit the covariance structures better than RR models [4, 5, 8]. However, they are less often used in genetic analyses due to the lack of user-friendly software for SAD models and to the difficulty in accounting for nonstationary longitudinal data without inflation of the number of parameters or without appropriate software for CP models.

When the correlation between several longitudinal traits needs to be taken into account, modeling the covariance structure of repeated records becomes even more complicated. Once again, RR models are more often used than CP or SAD models because their extension to multivariate situations is straightforward. However, such extensions can require a large number of parameters and the drawbacks described for univariate RR models remain [9]. Multiple-trait extension of the CP model is more complex and was discussed in Jaffrézic et al. [10]. The same authors also proposed an extension of the SAD model to accommodate multiple-trait situations [9]. Nonetheless, the use of these multiple-trait CP and SAD models is still challenging because of the lack of user-friendly and readily available software. The application of the multiple-trait RR, CP and SAD approaches is relevant for a wide range of traits that are repeatedly measured over time, such as traits related to milk production (milk yield, milk composition, and somatic cell counts [11]), animal growth (body weight and feed intake) or reproduction traits (litter size and birth weight [12]).

Breeding programs in polytocous species aim at increasing the number of young weaned per female. To reach this goal, they have primarily focused on litter size [13]. However, in some selection programs, response to selection using a simple repeatability model has been low for two reasons: litter sizes at different parities were not considered as different traits [14, 15], and larger litter sizes were correlated with decreased survival of the young [16], probably because of a smaller weight and a reduced level of maturity at birth [17].

The objective of our study was to propose a new multiple-trait SAD model (freely available software) that can take correlations among several longitudinal traits into account. To illustrate the functionality of our model, we used it to analyze litter size and average birth weight in two species, pigs and rabbits.

Methods

Multiple-trait SAD models

Let $y_i(t_j)$ be the observation of animal i at time t_j . All linear mixed models used to study repeated measures of y_i over time can be decomposed as follows:

$$y_i(t_j) = \mu_i(t_j) + u_i(t_j) + p_i(t_j),$$

where $\mu_i(t_j)$ represents the fixed effects at time t_j , and $u_i(t_j)$ and $p_i(t_j)$ the genetic and pseudo-permanent animal effect random functions, with covariance functions $\mathbf{U}(t_j, t_{j'})$ and $\mathbf{P}(t_j, t_{j'})$, respectively. Note that this model does not include a residual term in order to help convergence and avoid identifiability problems between structured permanent and classical residual covariance matrices [18]. Thus, the residual variance was, by definition, included in the (co)variance matrix of the pseudo-permanent effects included in the model. In the single-trait situation, possible non-null covariance between random effects at different times [i.e. $\mathbf{U}(t_j, t_{j'})$ and $\mathbf{P}(t_j, t_{j'})$] are taken into account in the SAD approach by modeling the form of the random-effects functions. Specifically, it assumes that a random effect at time t_j can be explained by the previous random effects (i.e. at time $t_k, k < j$). For instance, for a given random effect $\mathbf{p}(t)$, the general form of the SAD model of order α is:

$$\mathbf{p}(t_j) = \sum_{s=1}^{\alpha} \theta_{sj} \mathbf{p}(t_{j-s}) + \mathbf{e}(t_j),$$

where θ_{sj} is the s th antedependence parameter for time t_j , and $\mathbf{e}(t_j)$ is a random normally distributed effect (error term) with mean 0 and innovation variance $\sigma_p^2(t_j)$. To reduce the number of parameters in the SAD model, θ_{sj} and $\sigma_p^2(t_j)$ are assumed to be continuous functions of time: $\theta_{sj} = \sum_{q=0}^{\beta_s} a_{sq} t_j^q$ for a function of degree β_s and $\sigma_p^2(t_j) = \exp\left(\sum_{q=0}^{\gamma} b_q t_j^q\right)$ for a function of degree γ . A single-trait SAD model is then defined by the order of the antedependence (α), the degree of the polynomial for each antedependence parameter (β_1 to β_α), and the degree of the polynomial for the innovation variance (γ) for each random effect. We will refer to single-trait SAD models as SAD $\alpha\beta_1 \dots \beta_\alpha\gamma$ [8]. For instance SAD 111 stands for a SAD model with:

$$\mathbf{p}(t_j) = \theta_{1j} \mathbf{p}(t_{j-1}) + \mathbf{e}(t_j), \quad \theta_{1j} = a_{1,0} + a_{1,1} t_j$$

$$\text{and } \sigma_p^2(t_j) = \exp(b_0 + b_1 t_j).$$

We propose an extension of the single-trait SAD model to the multiple-trait situation by assuming that, in addition to the within-trait antedependence relationship, a random effect of one trait can be a function of the same random effect of the other traits considered. For two traits \mathbf{y}_1 and \mathbf{y}_2 , the general form of the multiple-trait SAD model of order $\alpha, \alpha', \eta, \eta'$, for a given random effect \mathbf{p} , can be written as (for $j > \max(\alpha, \alpha', \eta, \eta')$):

$$\begin{aligned}
 \mathbf{p}_1(t_j) &= \sum_{s=1}^{\alpha} \theta_{sj} \mathbf{p}_1(t_{j-s}) + \sum_{s=c}^{\eta} \delta_{sj} \mathbf{p}_2(t_{j-s}) + \mathbf{e}_1(t_j), \\
 \mathbf{p}_2(t_j) &= \sum_{s=1}^{\alpha'} \theta'_{sj} \mathbf{p}_2(t_{j-s}) + \sum_{s=c'}^{\eta'} \delta'_{sj} \mathbf{p}_1(t_{j-s}) + \mathbf{e}_2(t_j), \quad (1)
 \end{aligned}$$

where θ_{sj} and θ'_{sj} are the s th antedependence parameters at time j for traits 1 and 2, respectively, and δ_{sj} and δ'_{sj} are the $(s - c + 1)$ th or $(s - c' + 1)$ th cross-antedependence parameters at time j for traits 1 and 2, respectively. Note that, in contrast to the antedependence relationship that starts at time t_{j-1} , the cross antedependence relationships show greater flexibility and start at time t_{j-c} ($t_{j-c'}$), with c (c') greater or equal to 0. Here, $\mathbf{e}_1(t_j)$ and $\mathbf{e}_2(t_j)$ are normally distributed random effects with mean 0 and innovation variance $\sigma_{p1}^2(t_j)$ and $\sigma_{p2}^2(t_j)$, respectively. Error terms \mathbf{e}_1 and \mathbf{e}_2 are assumed to be independent, except if $c > 0$ and $c' > 0$ when a correlation between the two can be considered at time t_1 . This constraint on the correlation between error terms ensures the identifiability of the parameters in the multiple-trait SAD model, as has been demonstrated for structural equation models (SEM). Indeed, the multiple-trait SAD model can be considered to be a specific kind of SEM [19] in which the structural parameters are functions of time. Consider, for the sake of simplicity, the simple case of no repetition per subject. A SEM with a recursive relationship between two traits \mathbf{y}_1 and \mathbf{y}_2 for animal i is:

$$\begin{cases} y_{1i} = \mathbf{x}'_{1i} \boldsymbol{\beta}_1 + u_{1i} + \varepsilon_{1i} \\ y_{2i} = \lambda y_{1i} + \mathbf{x}'_{2i} \boldsymbol{\beta}_2 + u_{2i} + \varepsilon_{2i} \end{cases} \Leftrightarrow \begin{cases} y_{1i} = \mathbf{x}'_{1i} \boldsymbol{\beta}_1 + u_{1i} + \varepsilon_{1i} \\ y_{2i} = \lambda \mathbf{x}'_{1i} \boldsymbol{\beta}_1 + \mathbf{x}'_{2i} \boldsymbol{\beta}_2 + \lambda u_{1i} + u_{2i} + \lambda \varepsilon_{1i} + \varepsilon_{2i} \end{cases} \quad (2)$$

Rosa et al. [20] showed that parameter identifiability in Eq. (2) is possible by assuming independency between residuals ε_1 and ε_2 . Varona et al. [12] proposed an extension of the SEM that allows for unequal recursive relationships between random terms:

$$\begin{cases} y_{1i} = \mathbf{x}'_{1i} \boldsymbol{\beta}_1 + u_{1i} + \varepsilon_{1i} \\ y_{2i} = \lambda \mathbf{x}'_{1i} \boldsymbol{\beta}_1 + \mathbf{x}'_{2i} \boldsymbol{\beta}_2 + \lambda_u u_{1i} + u_{2i} + \lambda_p \varepsilon_{1i} + \varepsilon_{2i} \end{cases} \quad (3)$$

Identifiability of the parameters in Eq. (3) is achieved by assuming independency between \mathbf{u}_1 and \mathbf{u}_2 and between ε_1 and ε_2 [12]. If we discard antedependency because no repetition occurs and focus on the cross-antedependency, model (3) [and (2)] is equivalent to a multiple-trait SAD model with $c > \eta$ (recursive but not simultaneous relationship), $c' = 0$ and $\eta' = 0$ and a polynomial cross-antedependence function of degree 0 for both the genetic and pseudo-permanent environmental effects. Terms \mathbf{u}_1 , \mathbf{u}_2 and ε_1 , ε_2 are equivalent to the error terms for the genetic and pseudo-permanent environmental effects of a multiple-trait

SAD model. Following recommendations for identifiability in the SEM, identifiability in the SAD model is achieved by assuming independency between the error terms in Eq. (1) for both genetic and pseudo-permanent environmental effects. In the more general case of repeated measurements over time for both traits, if $c > \eta$ and $c' = 1$ (recursive cross-antedependence) for random effects, the SAD model at time t_1 is equivalent to a classical multiple-trait model (where parameters are identifiable even if random effects are correlated between traits). Correlation between the error terms of random effects at time t_1 is thus permitted in the SAD model without adversely affecting parameter identifiability. Although the previous considered a simple situation with two traits, the same SAD models for each random effect, and a recursive relationship, it is straightforward to extend the reasoning to more complicated SAD models.

As for the single-trait model, (cross-)antedependence parameters and innovation variances were assumed to be continuous functions of time. The multiple-trait SAD model is then defined for two traits by the order of the antedependence for each trait (α, α'), the starting points (c, c'), the order of the cross-antedependence ($\eta - c + 1, \eta' - c' + 1$), the degree of the polynomial for each (cross-)antedependence parameter, and the degree of the polynomial for the innovation variance of each trait (γ, γ'), as well as an indicator of the presence of an initial correlation between $\mathbf{e}_1(t_1)$, $\mathbf{e}_2(t_1)$. An interesting computational property of this multiple-trait SAD model is that, as is the case for the single-trait model,

the inverse of the covariance matrix \mathbf{P} can be easily calculated by the following Cholesky decomposition [21]: $\mathbf{P}^{-1} = \mathbf{L}' \mathbf{D}^{-1} \mathbf{L}$, where \mathbf{D} is a diagonal matrix with innovation variances as components, \mathbf{L} is a lower triangular matrix with 1s on the diagonal and the negatives of the (cross-)antedependence parameters and the initial correlation (ρ) between \mathbf{e}_1 , \mathbf{e}_2 as diagonal entries. For instance, for two traits, three time points, all (cross-)antedependence of order 1, $c = c' = 1$ and an initial correlation ρ between $\mathbf{e}_1(t_1)$ and $\mathbf{e}_2(t_1)$, the \mathbf{L} matrix for the vector [$\mathbf{p}_{y1}(1)$ $\mathbf{p}_{y2}(1)$ $\mathbf{p}_{y1}(2)$ $\mathbf{p}_{y2}(2)$ $\mathbf{p}_{y1}(3)$ $\mathbf{p}_{y2}(3)$] is:

$$\mathbf{L} = \begin{bmatrix} 1 & & & & & \\ -\rho & 1 & & & & \\ -\theta_{12} & -\delta_{12} & 1 & & & \\ -\delta'_{12} & -\theta'_{12} & 0 & 1 & & \\ 0 & 0 & -\theta_{13} & -\delta_{13} & 1 & \\ 0 & 0 & -\delta'_{13} & -\theta'_{13} & 0 & 1 \end{bmatrix}.$$

This multiple-trait SAD model can be easily fitted to data using ASReml [22], as well as the OWN Fortran

program that we have developed (available online at <https://zenodo.org/record/192036#.WEAYLdLhBaQ>).

The orders of (cross-)antedependence and the degrees of the polynomial functions of time for (cross-)antedependence parameters and innovation variances can be selected by comparing nested models using the likelihood ratio test and by comparing non-nested models using the Bayesian or Akaike information criterion (AIC) [23, 24]. To reduce the number of models that need to be compared, we suggest selecting the specification (order and degree) of the antedependence and innovation variances by first using a single-trait SAD model and then the cross-antedependence specification by using a multiple-trait SAD model. For the single-trait SAD model step of the selection process, we suggest to start with a SAD 100, then test an increase of the degree of the function of time (γ) for the innovation variance (i.e. SAD 101), and finally test an increase of the degree of the function of time for the antedependence parameter (β_1) (i.e. SAD 111). This procedure is repeated until there is no additional significant improvement of the model. Then, an increase of the order of the dependence (α) can be tested, starting with a constant second order antedependence parameter. For instance, if the last model of order 1 selected is SAD 121, the next model tested is SAD 2201. The increase of the polynome of time for the second order antedependence parameter can then be tested (i.e. SAD 2211) and so on. Using this step-by-step selection procedure, all models are nested and can be compared using the likelihood ratio test.

Data application

The multiple-trait SAD model was applied to data from two polytocous species (pigs and rabbits) to study the relationship between the number of young born alive per litter (litter size, LS) and the average birth weight (ABW) per litter, calculated as the sum of the weight at birth of all the individuals of a litter divided by LS. The pig dataset included 4345 litters from 1817 Large White sows over a maximum of five successive farrowings. The mean LS was 12.1 ± 3.6 . Piglets were weighed at birth in all litters, the mean ABW was 1514 ± 31 g. The rabbit dataset included 8706 litters from 2286 L-1777 does [25] over a maximum of five successive kindlings. Kittens were weighed at birth for 3490 litters. The mean LS and ABW were 9.5 ± 3.1 and 83 ± 14 g, respectively. The descriptive statistics of LS and ABW per parity are in Table 1.

Fixed effects included in the model were initially selected separately for each trait using a step-by-step descending procedure. To do this, simple models that did not take relationships between animals into account were applied to the data, nested models were then compared using the likelihood ratio test. The same within-species

Table 1 Litter size and average birth weight by parity in pigs and rabbits

Parity	Number of litters	LS (s.e.)	Number of litters	Average BW (s.e.)
<i>Pigs</i>				
1	1508	11.7 (3.4)	1508	1419 (29)
2	1134	12.0 (3.6)	1134	1565 (31)
3	861	12.6 (3.6)	861	1564 (30)
4	515	12.5 (3.5)	515	1566 (31)
5	327	12.4 (3.3)	327	1566 (31)
<i>Rabbits</i>				
1	2013	8.1 (2.7)	115	84 (17)
2	1868	9.5 (3.0)	955	85 (15)
3	1812	10.0 (3.2)	1174	83 (14)
4	1638	10.3 (3.2)	907	82 (13)
5	1375	10.0 (3.0)	339	82 (11)

fixed effects were included in the genetic multiple-trait SAD model for both traits: parity (five classes), proportion of females in the litter (covariate), and the combination of year and month of delivery (121 levels) for rabbits; and parity (five classes), season of farrowing (four classes), sire breed (seven classes) and sow weight when entering the farrowing unit (covariates) for pigs. It should be noted that the contemporary group effect (constant over parities) was included in the model as a random effect for pig data, in addition to the genetic and pseudo-permanent environmental effects.

ABW and LS were defined as traits of the sow/doe and analyzed using the previously described multiple-trait SAD model with successive time points at each farrowing/kindling. Data were also analyzed using a multiple-trait RR model in order to provide a comparison of the SAD multiple-trait model with the most currently used method. Selection of the degree of the Legendre polynomials for the permanent and genetic effects was performed for each trait by comparing nested single-trait RR models using the log likelihood ratio test. Then, multiple-trait RR models were applied to the data using the selected degree of polynomials for each trait. Goodness-of-fit of the SAD and RR models to the data were compared using the AIC.

Results

In rabbits, model selection for the antedependence relationship using single-trait analysis showed that the most appropriate SAD models were SAD 111 for the animal genetic effect on both traits (LS and ABW), and SAD 100 and SAD 111 for the pseudo-permanent environmental effect on LS and ABW, respectively. These antedependence characteristics were retained for the multiple-trait

model. Using the multiple-trait SAD model, the same cross-antependence characteristics were selected for the genetic and permanent environmental effects, that is a recursive relationship (i.e. $c > \eta$, in other words, LS random effects are not functions of ABW random effects), $c' = \eta' = 0$ (i.e. ABW random effects at parity j are functions of LS random effects at parity j) and a degree 2 for the cross-antependence function of time, which resulted in 20 parameters to model **U** and **P**. Let $LS_i(j)$ and $ABW_i(j)$ be the LS and ABW of dam i at parity j , the selected multiple-trait SAD model then is:

$$LS_i(j) = \mathbf{x}'_{ij}\boldsymbol{\beta}_{LS} + u_i(j) + p_{LS,i}(j),$$

$$ABW_i(j) = \mathbf{x}'_{ij}\boldsymbol{\beta}_{ABW} + v_i(j) + p_{ABW,i}(j), \text{ with}$$

$$u_i(j) = \theta_{u,1j}u_i(j-1) + e_{u,i}(j),$$

$$v_i(j) = \theta'_{v,1j}v_i(j-1) + \delta'_{v,0j}u_i(j) + e_{v,i}(j), \text{ and}$$

$$p_{LS,i}(j) = \theta_{p,1j}p_{LS,i}(j-1) + e_{p_{LS,i}}(j),$$

$$p_{ABW,i}(j) = \theta'_{p,1j}p_{ABW,i}(j-1) + \delta'_{p,0j}p_{LS,i}(j) + e_{p_{ABW,i}}(j).$$

In pigs, the SAD 101 model was retained for the within-trait antedependency for the genetic and pseudo-permanent environmental effects for LS, and SAD 111 and SAD 101 for the genetic and pseudo-permanent environmental effects for ABW, respectively. Recursive cross-antependence functions of order 1 with $c' = 0$, degree 0 and 1 were selected for the genetic and pseudo-permanent environmental effects, respectively. The random contemporary group effects of the two traits were independent.

After model selection, the single trait RR models selected in pigs included a Legendre polynomial of degree 1 for the genetic effects and a constant permanent effect over time for both traits. In rabbits, the best single trait RR model was of degree 1 for both the genetic and permanent environmental effects for both traits. Unfortunately, the multiple-trait extension of these models did not converge in rabbits. Thus, the multiple-trait RR model considered for rabbits consisted, as for pigs, of a polynomial function of degree 1 for genetic effects only.

In pigs, the AIC for the multiple-trait SAD and RR models were 66,319 and 66,326, respectively. For rabbits, the AIC were 70,684 and 70,911 for the multiple-trait SAD and RR models, respectively.

Heritability estimates with the SAD model for the two traits are in Table 2. Heritability estimates were moderate for all traits at all parities, with the exception of LS

Table 2 Heritability estimates obtained with the SAD model in pigs and rabbits

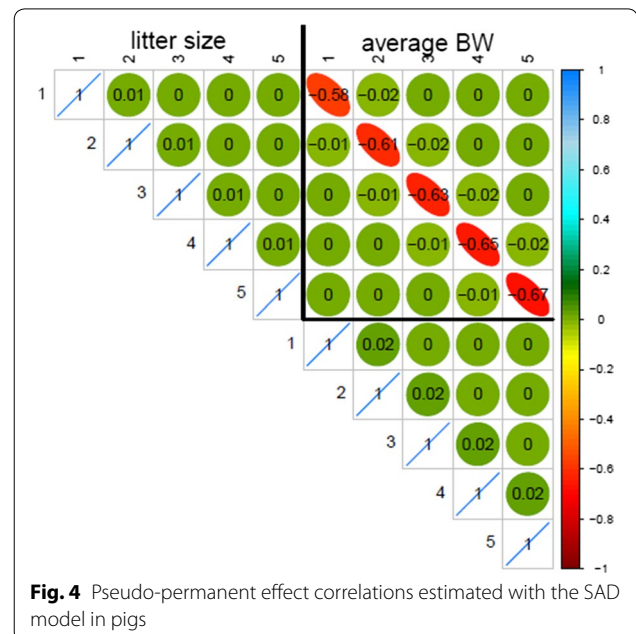
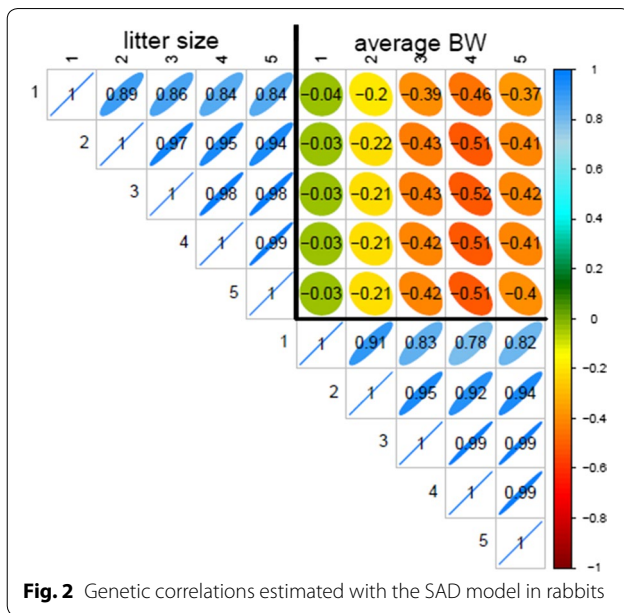
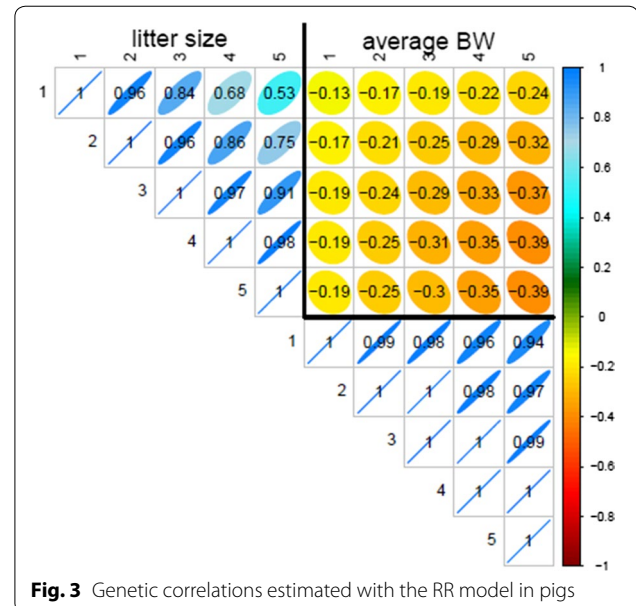
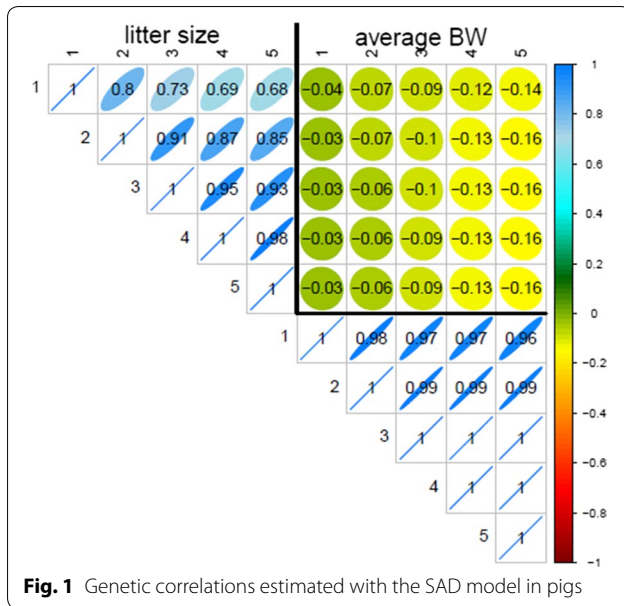
Parity	Pigs		Rabbits	
	Litter size	Average BW	Litter size	Average BW
1	0.21	0.29	0.09	0.23
2	0.25	0.33	0.19	0.26
3	0.24	0.35	0.27	0.32
4	0.22	0.33	0.29	0.37
5	0.19	0.30	0.24	0.39

at parity 1 in rabbits, which had a lower heritability estimate, i.e. 0.09. Heritability estimates for ABW were generally higher than for LS. In pigs, heritability estimates were quite stable across parities, ranging from 0.19 to 0.25 for LS and from 0.29 to 0.35 for ABW. In rabbits, heritability estimates tended to increase with parity for both traits, ranging from 0.09 to 0.29 for LS and from 0.23 to 0.39 for ABW. In pigs, heritability estimates obtained with the RR model tended to be slightly lower than those obtained with the SAD model, ranging from 0.18 to 0.28 for LS and from 0.22 to 0.31 for ABW. In rabbits, compared to the SAD model, heritability estimates obtained with the RR model were slightly lower for LS (ranging from 0.08 to 0.25) and higher for ABW (ranging from 0.24 to 0.42).

Genetic correlation matrices estimated with the multiple-trait SAD models are shown in Figs. 1 and 2 for pigs and rabbits, respectively. Estimates of the genetic correlation between LS tended to decrease as the distance between parities increased, the decrease being slightly more pronounced for pigs than for rabbits (0.84 versus 0.68 between parities 1 and 5). The same general pattern, i.e. a decrease of the estimated genetic correlation between more distant parities, was observed for ABW in rabbits. In pigs, estimates of the genetic correlations between ABW at different parities were high, the minimum being 0.96 between parities 1 and 5. Estimates of the genetic correlation between LS and ABW were negative in both species. In general, we observed an increase of the genetic antagonism between the two traits with the parity level for ABW regardless of the parity level for LS.

The genetic correlation matrix estimated with the RR model for pigs is in Fig. 3. The general pattern of these correlation estimates was similar to the one obtained with the SAD model for both species but with a more pronounced genetic antagonism between the two traits in pigs and less pronounced in rabbits (result not shown, the lowest correlation between traits in rabbits was equal to -0.34).

Correlation matrices for the pseudo-permanent environmental effects obtained with the multiple-trait SAD

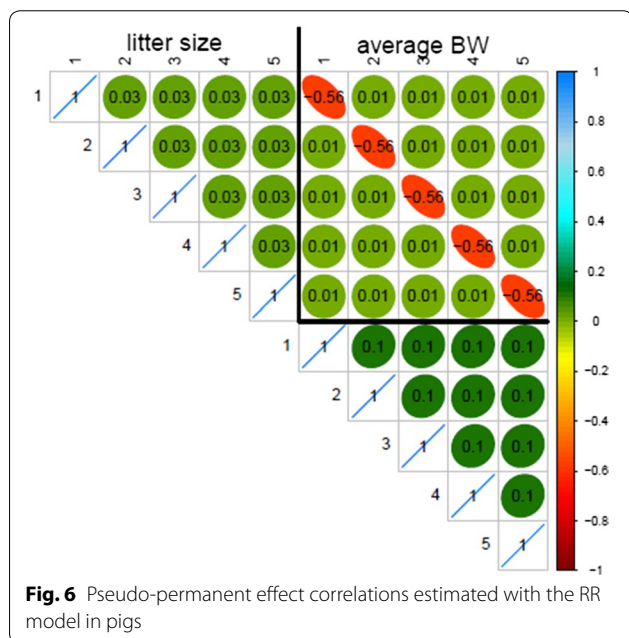
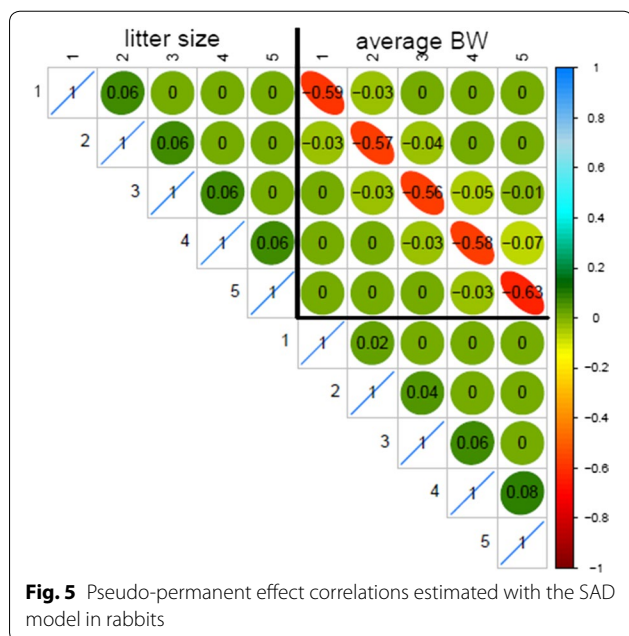


model are in Figs. 4 and 5 for pigs and rabbits, respectively. Estimates of the correlations between pseudo-permanent environmental effects at different parities within-trait were close to 0. The same across-trait pattern of estimates of correlations between pseudo-permanent environmental effects of the two traits for the same parity level (ranging from -0.57 to -0.67) and correlations close to 0 otherwise. The same general pattern of

correlations was obtained with the RR model for pigs (Fig. 6) and rabbits (result not shown).

Discussion

We propose in this article a multiple-trait SAD model that takes the within-trait and cross-trait correlations over time into account. A first multiple-trait extension of the SAD model was proposed a few years back by Jaf-frézic et al. [9], however our multiple-trait SAD model



differs from theirs regarding the assumption made about the innovation covariance matrix. For their extension, Jaffrézic et al. [9] assumed that the error terms $e_1(t)$ and $e_2(t)$ are correlated as a function of time. To insure identifiability of the parameters, our assumption is that these error terms are independent, except at time t_1 , when they can be correlated if $c > 0$ and $c' > 0$ in Eq. (1).

We suggest a two-step procedure for selecting the order of (cross-)antependence and degree of the polynomial

functions: within-trait selection of the antependence parameters using a single-trait SAD model, followed by selection of the cross-antependence parameters using the multiple-trait SAD model. This same two-step approach was performed by Jaffrézic et al. [9]. In the multiple-trait SAD model, selection of the order of cross-antependence may differ according to the goal of the analysis. If the goal is to estimate the correlation between traits, our experience showed that a recursive cross-antependence with zero as starting point and order 1 (i.e. $c' = \eta' = 0$) is generally sufficient to model all forms of correlation across traits. In such cases, only the degree of the cross-antependence function needs to be selected by comparing nested models using the likelihood ratio test. Conversely, if the goal of the analysis is to study causal relationships between traits, it seems necessary to define, a priori, the general pathway between traits, as in SEM (resursiveness, simultaneity), to fit the order of the cross-antependence in the multiple-trait SAD model to match the pathway. Then, the degree of the polynomial function for each cross-antependence parameter is selected by comparing nested models using the likelihood ratio test. In such situations, antependence parameter values are of interest in addition to the \mathbf{U} and \mathbf{P} covariance matrices.

Similar to a RR model, the number of breeding values predicted per animal with a SAD model equals the number of time points. Eigen decomposition (eigenvalues, eigenvectors) of the genetic covariance matrix obtained with the SAD model can help to determine which linear combination of breeding values can be used for genetic selection purposes. Nonetheless, given the way the SAD covariance matrix is structured, a decomposition in eigenfunctions (continuous function of time) similar to that proposed for RR models [26] is certainly not feasible.

We applied the multiple-trait SAD model to study the correlation between LS and ABW over parities in two species. For this purpose, we used a recursive cross-antependence relationship. To study the causal relationship between these two traits with a SEM, Varona et al. [12] proposed a one-way causal path that established an effect of LS on ABW within parity as the most likely general pathway between LS and ABW. If our goal was to study the causal relationship between traits, similar to Varona et al. [12], we would have needed to add cross-antependence parameters to the SAD model that links the random parameters of ABW at time t to those of LS at time $t + 1$. As in previous studies, we considered LS and ABW as traits of the dam [12, 13].

Heritability estimates for LS were higher than the values of about 0.10 that were reported in previous studies [27–31] but were consistent with the total heritability for LS reported for the same pig breed by Kaufmann et al.

(0.24) [32] and for rabbits by Nagy et al. (0.11–0.31) [33]. The higher heritabilities obtained in the current study can be explained by the fact that heritabilities estimated with a model that assumes different traits over time are generally higher than those obtained by using a simple repeatability model [34]. An increase of the heritability of LS with parity similar to that observed here was previously reported for both species [28–30], however it was not observed in pigs by Lukovic et al. [35] using records over 10 parities. Our heritability estimates for ABW were consistent with those reported by Hermesch in pigs (0.31) [13] but higher than those reported by Varona et al. in pigs (0.23) [12] and Bolet et al. [36] (0.04) or Garreau et al. [37] (0.06) in rabbits.

On the one hand, the decrease in the within-trait genetic correlation with distance between measurements (parities) is a result frequently reported in the literature [38, 39] and the genetic correlation matrix reported by Hanenberg et al. [28] for LS in Dutch Landrace pigs is close to the estimates obtained in our study. On the other hand, estimates of the genetic correlation for ABW reported for Australian pigs tended to be lower than obtained in the current study, but their standard errors were high [40]. It is generally recommended to consider LS and ABW at first parity as different traits from the performances at later parities [28, 40]. The same conclusion can be drawn from our results, except for ABW in pigs, which can be considered as a repeatable trait, the lowest genetic correlation value being 0.95 between parity 1 and 5.

Regarding cross-correlations, our results showed that the general trend was an unfavorable genetic correlation between LS and ABW (slightly negative in pigs and more strongly negative in rabbits). Several studies have reported the same negative genetic correlation between these traits in pigs [31, 40, 41]. Even if the trend was less clear in pigs, we observed the same cross-correlation tendency over parities: ABW tended to be more negatively correlated to LS in late parities than in first parity. In other words, if one considered that LS occurs before ABW and studies the relationship between successive traits (i.e. ABW at time $t - 1$ has an “effect” on LS at time t which has an “effect” on ABW at time t and so on), our results show that the antagonism between successive traits increases with parity. This could be considered as an increasing adaptation of traits to their environmental conditions with time, LS at time t being part of the environment for ABW at time t , which in turn is the environment to which the animal has to adapt its LS at time $t + 1$ and so on. Hermesch et al. [40] also studied cross-correlations over parities in pigs and did not find any trend, but their estimates had large standard errors.

The close to 0 correlations that we found between pseudo-permanent effects of the same trait were previously reported between first and second parities in rabbits [30] and pigs [40]. Consistent with our results, Hermesch et al. [40] reported a strong negative correlation between the pseudo-permanent effects of the two traits at the same parity and no correlation between traits at different parities. In our model, the so-called pseudo-permanent environmental effect combines the dam characteristics that are not under genetic control and persist over parities and the environmental factors that are not taken into account in the model because not measured/observable. The correlation between pseudo-permanent environmental effects at different parities was extremely low. Thus, the pseudo-permanent effects probably mainly reflect specific animal parity-related factors that are not taken into account in the model or environmental factors that affect traits differently during the reproductive period. Given the high negative correlation between the pseudo-permanent environmental effects between traits, these unobservable factors have opposite effects on LS and ABW.

We compared results obtained with the multiple-trait SAD and RR models. The latter is the approach most often used for analyzing longitudinal data in genetic studies. Probably due to the small number of repetitions, the best RR model selected in pigs and the multiple-trait RR model that converged in rabbits only considered a Legendre polynomial of degree 1 for genetic effects. Thus, with the same number of parameters in pigs (16), the RR model was less flexible than the multiple-trait SAD model to model the variance covariance matrix of environmental effects. Comparison of AIC values showed that the multiple-trait SAD model provided a better fit to the data than the RR model in both species. The same conclusion has been drawn in previous studies on other traits [8, 9]. Results obtained with RR models were essentially consistent with those of the multiple-trait SAD model, which enables us to be confident in the estimations obtained with the SAD approach.

Conclusions

In this paper, we have outlined a multiple-trait SAD model to simultaneously analyze repeated measurements of several traits. This flexible approach was developed to provide an advantageous approach to model the covariance matrix of random terms with few parameters. When used to study relationship between LS and ABW over five parities in two species, the multiple-trait SAD model showed that a total of 16 or 20 parameters was sufficient to model the random term covariance matrices. This is much less than the number of parameters required to

obtain estimates when unstructured covariance matrices are assumed for \mathbf{U} and \mathbf{P} (up to 110 parameters). Furthermore, we showed that multiple-trait SAD models provide a better fit to the data than multiple-trait RR models. We offer a freely-available online Fortran program that can be used to implement this SAD model in ASReml.

Authors' contributions

YB and EB provided pig and rabbit data, respectively. HG and LC prepared the pig and rabbit data, respectively. They all took part in the interpretation of the results. ID developed the multiple-trait SAD model and wrote the Fortran program, performed data analysis and wrote the paper. All authors read and approved the final manuscript.

Author details

¹ GenPhySE, INRA, INPT, ENVT, Université de Toulouse, 31326 Castanet-Tolosan, France. ² Pectoul, INRA, 31326 Castanet-Tolosan, France. ³ GenESI, INRA, 17700 Surgères, France.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The Fortran programs for the OWN function are freely available at <https://zenodo.org/record/192036#WEAYLdLhBaQ>. The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Received: 12 October 2016 Accepted: 10 January 2017

Published online: 20 January 2017

References

- Schnyder U, Hofer A, Labroue F, Kunzi N. Genetic parameters of a random regression model for daily feed intake of performance tested French Landrace and Large White growing pigs. *Genet Sel Evol*. 2001;33:635–58.
- Lorenzo Bermejo J, Roehe R, Schulze V, Rave G, Looft H, Kalm E. Random regression to model genetically the longitudinal data of daily feed intake in growing pigs. *Livest Prod Sci*. 2003;82:89–99.
- Manzanilla Pech CIV, Veerkamp RF, Calus MPL, Zom R, van Knegsel A, Pryce JE, et al. Genetic parameters across lactation for feed intake, fat- and protein-corrected milk, and liveweight in first-parity Holstein cattle. *J Dairy Sci*. 2014;97:5851–62.
- Jaffrézic F, Venot E, Laloe D, Vinet A, Renand G. Use of structured antedependence models for the genetic analysis of growth curves. *J Anim Sci*. 2004;82:3465–73.
- Jaffrézic F, Pletcher SD. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics*. 2000;156:913–22.
- Schaeffer L. Application of random regression models in animal breeding. *Livest Prod Sci*. 2004;86:35–45.
- Druet T, Jaffrézic F, Boichard D, Ducrocq V. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of french Holstein cows. *J Dairy Sci*. 2003;86:2480–90.
- David I, Ruesche J, Drouilhet L, Garreau H, Gilbert H. Genetic modeling of feed intake. *J Anim Sci*. 2015;93:965–77.
- Jaffrézic F, Thompson R, Hill WG. Structured antedependence models for genetic analysis of repeated measures on multiple quantitative traits. *Genet Res*. 2003;82:55–65.
- Jaffrézic F, Thompson R, Pletcher SD. Multivariate character process models for the analysis of two or more correlated function-valued traits. *Genetics*. 2004;168:477–87.
- Jamrozik J, Bohmanova J, Schaeffer L. Relationships between milk yield and somatic cell score in Canadian Holsteins from simultaneous and recursive random regression models. *J Dairy Sci*. 2010;93:1216–33.
- Varona L, Sorensen D, Thompson R. Analysis of litter size and average litter weight in pigs using a recursive model. *Genetics*. 2007;177:1791–9.
- Hermesch S. Avenues for genetic improvement of litter size and litter mortality. In: *Proceedings of the AGBU Pig Genetics Workshop: 7–8 November 2001; Armidale; 2001*. p. 36–43.
- Irgang R, Fávero JA, Kennedy BW. Genetic parameters for litter size of different parities in Duroc, Landrace, and large white sows. *J Anim Sci*. 1994;72:2237–46.
- Baselga M, Gomez E, Cifre P, Camacho J. Genetic diversity of litter size traits between parities in rabbits. *J Appl Rabbit Res*. 1992;15:198–205.
- Johnson RK, Nielsen MK, Casey DS. Responses in ovulation rate, embryonal survival, and litter traits in swine to 14 generations of selection to increase litter size. *J Anim Sci*. 1999;77:541–57.
- Canario L, Père M, Tribout T, Thomas F, David C, Gogué J, et al. Estimation of genetic trends from 1977 to 1998 of body composition and physiological state of Large White pigs at birth. *Animal*. 2007;1:1409–13.
- Wang W. Identifiability of linear mixed effects models. *Electron J Stat*. 2013;7:244–63.
- Gianola D, Sorensen D. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics*. 2004;167:1407–24.
- Rosa GJM, Valente BD, de los Campos G, Wu X-L, Gianola D, Silva MA. Inferring causal phenotype networks using structural equation models. *Genet Sel Evol*. 2011;43:6.
- Pourahmadi M. Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation. *Biometrika*. 1999;86:677–90.
- Gilmour AR, Gogel BJ, Cullis BR, Thompson R. *ASReml User guide Release 3.01*, Hemel Hempstead: VSN International Ltd; 2009.
- Schwarz GE. Estimating the dimension of a model. *Ann Stat*. 1978;6:461–4.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control*. 1974;19:716–23.
- Garreau H, Duzert R, Tudela F, Baillot C, Ruesche J, Grauby G, et al. Gestion et sélection de la souche INRA 1777: résultats de trois générations de sélection. In: *Proceedings of Les 11^{èmes} Journées de la Recherche Cunicole: 29–30 November 2005; Paris; 2005*. p. 19–22.
- Van Der Werf JHJ, Goddard ME, Meyer K. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. *J Dairy Sci*. 1998;81:3300–8.
- Putz A, Tiezzi F, Maltecca C, Gray K, Knauer M. Variance component estimates for alternative litter size traits in swine. *J Anim Sci*. 2015;93:5153–63.
- Hanenberg EHAT, Knol EF, Merks JWM. Estimates of genetic parameters for reproduction traits at different parities in Dutch Landrace pigs. *Livest Prod Sci*. 2001;69:179–86.
- Roehe R, Kennedy BW. Estimation of genetic parameters for litter size in Canadian Yorkshire and Landrace swine with each parity of farrowing treated as a different trait. *J Anim Sci*. 1995;73:2959–70.
- Piles M, García M, Rafel O, Ramon J, Baselga M. Genetics of litter size in three maternal lines of rabbits: repeatability versus multiple-trait models. *J Anim Sci*. 2006;84:2309–15.
- Mérou I, Bernard E, Bidanel JP, Canario L. Genetic parameters for litter traits including farrowing duration and piglet survival up to weaning in the French Large White and Landrace sows. In: *Proceedings of the 16th Annual Meeting of the European Association of Animal Production: 23–27 August 2010; Heraklios*. 2010.
- Kaufmann D, Hofer A, Bidanel JP, Kunzi N. Genetic parameters for individual birth and weaning weight and for litter size of Large White pigs. *J Anim Breed Genet*. 2000;117:121–8.
- Nagy I, Radnai I, Nagyné-Kiszlinger H, Farkas J, Szendrő Z. Genetic parameters and genetic trends of reproduction traits in synthetic Pannon rabbits using repeatability and multi-trait animal models. *Arch Tierz*. 2011;54:297–307.
- Mekkawy W, Roehe R, Lewis R, Davies M, Bünger L, Simm G, et al. Comparison of repeatability and multiple trait threshold models for litter size in sheep using observed and simulated data in Bayesian analyses. *J Anim Breed Genet*. 2010;127:261–71.
- Luković Z, Uremović M, Konjačić M, Uremović Z, Vincek D. Genetic parameters for litter size in pigs using a random regression model. *Asian Aust J Anim Sci*. 2007;20:160–5.
- Bolet G, Zerrouki N, Gacem M, Brun JM, Lebas F. Genetic parameters and trends for litter and growth traits in a synthetic line of rabbits created in

- Algeria. In: Proceedings of the 10th World Rabbit Congress: 3–6 August 2012; Sharm El-Sheikh. 2012.
37. Garreau H, Bolet G, Larzul C, Robert Granié C, Saleil G, San Cristobal M, et al. Results of four generations of a canalising selection for rabbit birth weight. *Livest Sci*. 2008;119:55–62.
 38. Sawalha RM, Keown JF, Kachman SD, Van Vleck LD. Evaluation of autoregressive covariance structures for test-day records of Holstein cows: estimates of parameters. *J Dairy Sci*. 2005;88:2632–42.
 39. Nunez-Anton V, Zimmerman DL. Modeling non-stationary longitudinal data. *Biometrics*. 2000;56:699–705.
 40. Hermesch S, Luxford BG, Graser HU. Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs-3. Genetic parameters for reproduction traits and genetic correlations with production, carcase and meat quality traits. *Livest Prod Sci*. 2000;65:261–70.
 41. Huby M, Canario L, Tribout T, Caritez J, Billon Y, Gogué J, et al. Genetic correlations between litter size and weights, piglet weight variability and piglet survival from birth to weaning in Large White pigs. In: Proceedings of the 54th Annual Meeting of the European Association for Animal Production: 31 August–3 September 2003; Roma. 2003.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

