Published online October 5, 2017

Genetic structured antedependence and random regression models applied to the longitudinal feed conversion ratio in growing Large White pigs[1]

V. H. Huynh-Tran,[2] H. Gilbert, and I. David

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

ABSTRACT: The objective of the present study was to compare a random regression model, usually used in genetic analyses of longitudinal data, with the structured antedependence (SAD) model to study the longitudinal feed conversion ratio (FCR) in growing Large White pigs and to propose criteria for animal selection when used for genetic evaluation. The study was based on data from 11,790 weekly FCR measures collected on 1,186 Large White male growing pigs. Random regression (RR) using orthogonal polynomial Legendre and SAD models was used to estimate genetic parameters and predict FCR-based EBV for each of the 10 wk of the test. The results demonstrated that the best SAD model (1 order of antedependence of degree 2 and a polynomial of degree 2 for the innovation variance for the genetic and permanent environmental effects, i.e., 12 parameters) provided a better fit for the data than RR with a quadratic function for the genetic and permanent environmental effects (13 parameters), with Bayesian information criteria values of −10,060 and −9,838, respectively. Heritabilities with the SAD model were higher than those of RR over the first 7 wk of the test. Genetic correlations between weeks were higher than 0.68 for short intervals between weeks and decreased to 0.08 for the SAD model and −0.39 for RR for the longest intervals. These differences in genetic parameters showed that, contrary to the RR approach, the SAD model does not suffer from border effect problems and can handle genetic correlations that tend to 0. Summarized breeding values were proposed for each approach as linear combinations of the individual weekly EBV weighted by the coefficients of the first or second eigenvector computed from the genetic covariance matrix of the additive genetic effects. These summarized breeding values isolated EBV trajectories over time, capturing either the average general value or the slope of the trajectory. Finally, applying the SAD model over a reduced period of time suggested that similar selection choices would result from the use of the records from the first 8 wk of the test. To conclude, the SAD model performed well for the genetic evaluation of longitudinal phenotypes.

Key words: feed efficiency, longitudinal data, pigs, random regression, selection criterion, structured antedependence model

J. Anim. Sci. 2017.95
doi:10.2527/jas2017.1864

INTRODUCTION

Feed efficiency is a benchmark for profitability in pig farming because the cost of feed represents about two-thirds of total production costs. It also reduces the negative effects of livestock farming on the environment (Patience et al., 2015; Gilbert et al., 2017). With the development of automatic equipment, individual feed intake (**FI**) and BW values can be repeatedly measured during the production period in some species. The analysis of repeated records can provide more accurate estimations in a genetic selection context than simple trait analyses (Boligon et al., 2011). To analyze such longitudinal data, genetic models

[1]Our work is part of the H2020 Feed-a-Gene project funded by the European Union's H2020 Programme under grant agreement number 633531. The authors would like to thank the Animal Genetics Division at INRA, France, and the European project H2020 Feed-a-Gene for financial support for Huynh-Tran's PhD thesis. The staff of the experimental facilities GenESI (UE1372, Surgères, France) is acknowledged for breeding the animals and for data collection.

[2]Corresponding author: van-hung.tran@inra.fr

Received June 27, 2017.
Accepted August 28, 2017.

should account for the covariance structures of the repeated records with few parameters to estimate. The random regression (**RR**) model is widely used, even if it presents various drawbacks such as higher variances at the beginning and the end of the studied period, so-called border effect problems (Jaffrézic et al., 2004; Meyer, 2005). The structured antedependence (**SAD**) model also deals with the correlation structure of data and has been shown to better fit covariance structures than RR models (Jaffrézic and Pletcher, 2000; Jaffrézic et al., 2004; David et al., 2015). Up to now, it has been less widely used than the RR model due to the lack of tools; however, user-friendly software (David et al., 2017) is now freely available (https://zenodo.org/record/896377; accessed 20 Sep. 2017). For selection purposes, an interpretable eigenvalue decomposition of the additive genetic matrix **K** of the RR coefficients of the RR model has been proposed to summarize the individual genetic potential over time as 1 or 2 values (Van Der Werf et al., 1998), capturing features such as persistency or area under the curve when applied to lactation curves (Togashi and Lin, 2006). To our knowledge, no methods for summarizing breeding values from the SAD model have yet been proposed. The objective of our study was to compare RR and SAD models for the genetic analysis of repeated measures of the feed conversion ratio (**FCR**) in growing pigs and to propose criteria for animal selection for the SAD model.

MATERIAL AND METHODS

Data were collected in accordance with the applicable national regulations on livestock welfare in France.

Pigs and Data Collection

The present study includes data from 1,186 Large White boars over 8 generations of divergent selection for residual FI raised after weaning in the Rouillé INRA experimental farm (GenESI, Vienne, France). The selection process was described in detail by Gilbert et al. (2007). The data used were collected from candidate boars tested in groups of 12 in pens equipped with single electronic feeders (ACEMA 64; Skiold Acemo, Pontivy, France). Pigs were age 67 ± 1 d (25 ± 4 kg) at the beginning of the test and were tested during the growing–finishing period up to 168 ± 13 d (115 ± 11 kg). The records collected during the first week of the test, when pigs acclimated to the feeders, were discarded from the analysis. Animals were fed ad libitum with a pelleted diet of cereals and soybean meal with 10 MJ NE/kg and 160 g CP/kg, and a minimum of 0.80 g digestible Lys/MJ NE.

During the 14 consecutive weeks (from wk 2 to 15) of the test period, animals were weighed weekly. The individual FI of each animal was automatically recorded each time it used the feeder. Weekly averages of the daily FI (**WDFI**) were then computed for each animal. The WDFI outlier values and WDFI for which more than 2 d of records were missing in a given week were removed from the analysis, as reported by David et al. (2015). The FCR was calculated for each animal i and week j ($j \in \{4, \ldots, 13\}$) as follows (Huynh-Tran et al., 2017):

$$FCR_{ij} = WDFI_{ij}/ADG_{ij},$$

in which $WDFI_{ij}$ is the WDFI of animal i for week j and ADG_{ij} is the ADG of animal i for week j ($j \in \{4, \ldots, 13\}$) estimated over a 4-wk period as follows:

$$ADG_{ij} = (BW_{ij+2} - BW_{ij-2})/(age_{ij+2} - age_{ij-2}),$$

in which BW_{ij} and age_{ij} are the BW and the age of animal i at week j, respectively. Only animals with at least 3 measures of FCR over the 10-wk period (wk 4 to 13) were retained for analysis. Extreme values of FCR (<0 and >4.5) were considered outliers and set as missing. The final data set comprised 11,790 weekly FCR values for 1,186 male growing pigs available from wk 4 to 13 of the test. For the sake of simplicity, we will denote $t_j \in \{1, \ldots, 10\}$ instead of $j \in \{4, \ldots, 13\}$ hereafter. A total of 3,986 animals was included in the pedigree.

Data Analysis

Estimations of Genetic Parameters. Repeated longitudinal FCR measurements were analyzed using the RR and SAD models. Both models can be written, for animal i at time t_j, as

$$FCR_{ij} = \mu_i(t_j) + u_i(t_j) + p_i(t_j) + \varepsilon_{ij}, \qquad [1]$$

in which $\mu_i(t_j)$ is the fixed effect at time t_j; $u_i(t_j)$ and $p_i(t_j)$ are the random genetic and permanent environmental animal effects functions with $\mathbf{u} \sim N(0, \mathbf{G} \otimes \mathbf{A})$ and $\mathbf{p} \sim N(0, \mathbf{P} \otimes \mathbf{I})$, in which **A** is the known relationship matrix; **I** the identity matrix; and **G** and **P** the covariance matrices between weekly measurements of FCR (of dimension 10 × 10) for genetic and permanent environmental effects, respectively. Finally, ε_{ij} is the random residual effect $\varepsilon \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$. The random functions were independent from one another.

In the RR model, for a given random effect $u_i(t_j)$, the general form of the random function of order m is $u_i(t_j) = \sum_{k=0}^{m} a_{ik}\varphi_k(t_j)$, in which a_{ik} is the $(k + 1)^{th}$ RR coefficient for the genetic effects for animal i, with $\mathbf{a} \sim$

$N(0,\mathbf{K}\otimes\mathbf{A})$, in which **K** is the covariance matrix of the additive RR coefficients, and $\varphi_k(t_j)$ is the $(k+1)^{th}$ Legendre polynomial at time t_j. In the RR model, the relationship between **K** and **G** is given by $\mathbf{G_{RR}} = \varphi\mathbf{K}\varphi'$, in which φ is the $n \times (m+1)$ (in which n is the number of time points) matrix of the Legendre polynomials for all time points.

In the SAD model, each random function is defined by 3 parameters: the order of the antedependance (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). The function for the random effect **u** is $u_i(t_j) = \sum_{s=1}^{\alpha} \theta_{sj} u_i(t_{j-s}) + e_i(t_j)$, in which θ_{sj} is the s^{th} antedependence parameter for time t_j and $e_i(t_j)$ is the error term for animal i at time t_j; **u** and **e** are independent and $e(t_j) \sim N(0, \mathbf{A}\sigma_e^2(t_j))$. To reduce the number of parameters in the SAD model, continuous functions of time were assumed for antedependence parameters $\theta_{sj} = \sum_{q=0}^{\beta_s} d_{sq} t_j^q$ and for the innovation variance $\sigma_e^2(t_j) = \exp\left(\sum_{q=0}^{\gamma} b_q t_j^q\right)$, in which d_{sq} and b_q are the coefficients for antedependence parameters and innovation variance.

We noted a SAD model with a given set of parameters as follows: SAD$\alpha - \beta_1, \ldots, \beta_\alpha\gamma$. To facilitate convergence and avoid identifiability problems between the structured random permanent environmental effect and the residual covariance matrices (Wang, 2013), the residual term ε_{ij} was removed from Eq. [1] for the SAD model. The residual variance is therefore, for this approach, included in the covariance matrix of the permanent environmental effect.

Covariance components were estimated for both models using the REML method using ASReml software (Gilmour et al., 2009). Estimations for SAD models were computed using the OWN function that allows users of ASReml to model their own variance structure, as proposed by David et al. (2017). The fixed effects included were the same for both models, as previously described by Huynh-Tran et al. (2017).

Both the degree of the polynomial functions for the RR approach and the order and degrees of the antedependence functions in the SAD approach were selected by comparing nested models using likelihood ratio tests. Once the best model for each approach was identified, the data-fitting capacity of the selected RR and SAD models was compared using the Bayesian information criteria (**BIC**; Schwarz, 1978): $BIC = -2\ln(L) + c \times \ln(N-p)$, in which L is the REML of the model, N is the number of observations, and p and c are the number of fixed effects and covariance parameters, respectively. The approach (RR or SAD) with the lowest BIC was considered the best fit for the data.

We compared the heritability and the estimated the genetic covariance matrices obtained using the 2 approaches. The heritability estimates were computed for each week j as $h_j^2 = \hat{\mathbf{G}}_{jj} / \left(\hat{\mathbf{G}}_{jj} + \hat{\mathbf{P}}_{jj} + \sigma_e^2\right)$, in which $\hat{\mathbf{P}}$ and $\hat{\mathbf{G}}$ are the estimates of matrices **P** and **G**, respectively; σ_e^2 is included for only the RR model. Standard errors of heritability estimates were calculated for the RR model in ASReml using the method proposed by Fischer et al. (2004). For the SAD model, analytical expressions of the SE are more difficult to obtain. Therefore, we used a bootstrap procedure to obtain SE for this model. The bootstrap steps were as follows:

1. In iteration l, sample a vector $\mathbf{v}_l$ of antedependence parameters and innovation covariance parameters (for instance, d_{s0} to $d_{s\beta_s}$, b_0 to b_γ are the parameters related to the genetic variance) using multivariate sampling $\mathbf{v}_l \sim \text{MVN}(\hat{\mathbf{v}}, \mathbf{V})$, in which $\hat{\mathbf{v}}$ is the vector of estimates, **V** is their covariance matrix estimated using ASReml, and MVN is multivariate normal distribution;
2. Using $\mathbf{v}_l$, calculate the genetic and permanent environmental variances for each time point (David et al., 2015) and then their heritabilities;
3. Repeat steps 1 and 2 10,000 times to obtain a vector of estimated heritabilities; and
4. Based on the vector of estimated heritabilities, calculate the mean and SE for the heritability (Efron and Hastie, 2016).

The EBV for each time point for the RR and SAD models were obtained as follows: EBV obtained with the SAD model (**EBV_SAD**) were provided in the ASReml outputs and the EBV obtained with the RR model (**EBV_RR**) were computed using the estimations of the individual regression coefficients provided by ASReml as $\text{EBV\_RR}_i(t_j) = \sum_{k=0}^{m} \hat{a}_{ik}\varphi_k(t_j)$. We denoted **sEBV_RR**$_i$ and **sEBV_SAD**$_i$ as the sum of the EBV_RR and EBV_SAD, respectively, for an animal i over the test period.

We compared these EBV with each other and with the overall breeding values (the EBV from the animal model using the FCR computed over the 10-wk period; **cEBV**) obtained by analyzing the FCR for the entire test period computed as the ratio of the ADFI during the 10 wk of test over the ADG for the same period. This overall FCR was analyzed using an animal mixed model: $FCR_i = \mu_i + u_i + \varepsilon_i$, in which FCR_i is the overall FCR for the entire test period for animal i, μ_i is the fixed effect, u_i is the animal additive genetic effect of animal i, and ε_i is the residual term.

Selection Criterion. Next, computations addressed the issue of defining for each model a criterion to select the best animals based on their 10 weekly

EBV values. First, the patterns of EBV variation over time were described using a trajectory classification approach that classified animals into different trajectory groups using a *k*-means approach with the Euclidean distance. This method used a hill-climbing algorithm jointly with expectation–maximization. The optimal number of clusters was chosen according to the Calinski–Harabatz criterion (Genolini et al., 2015).

Next, the information contained in the 10 EBV was summarized in a reduced number of n_{ind} independent variables using an eigendecomposition of the **G** matrices estimated using the RR and the SAD approaches. This method decomposes the covariance matrix into a set of independent eigenvectors and associated eigenvalues. Each eigenvalue represents the amount of variance explained by the associated eigenvector (Kirkpatrick et al., 1990). Summarized breeding values (**SBV**) associated to the *p*th eigenvector (**SBV*p*$_i$**) were calculated for each animal *i* by multiplying the coresponding eigenvector with the vector of EBV_i. The SBV were denoted **SBV_RR*p*** and SBV_SAD*p* when obtained from the *p*th eigenvector of the **G** matrices of the RR and SAD models, respectively.

For the RR model, we also calculated SBV obtained from the eigendecomposition of the **K** matrix as recommended for this model (Meyer and Hill, 1997; Van Der Werf et al., 1998). These summarized EBV were denoted **SBV_RRK*p***. The SBV_RRK*p* for animal *i* was given by $\text{SBV\_RRK}p_i = \sum_{l=0}^{m} \hat{a}_{il} k_{pl}$, in which k_{pl} is the l^{th} element of the p^{th} eigenvector of **K** ($p = 0, \ldots, m$). The eigendecomposition of the **K** matrix instead of the **G** matrix has the advantage of producing SBV that can be interpreted in regard to their variation over time. Actually, eigenfunctions of time can be obtained by multiplying the eigenvectors of **K** with the Legendre polynomials (Schnyder et al., 2001; Englishby et al., 2016). The **K** matrix is also usually of reduced dimension compared with **G**.

We then characterized, within each approach, the connection between the EBV group trajectory and the different SBV. In addition, to validate the interpretation of the SBV obtained with the 2 **G** matrices, they were compared with the SBV_RRK*p*. Finally, we compared these SBV with cEBV (EBV for the full test period FCR), sEBV_RR, and sEBV_SAD.

Appropriate Period for Estimating Longitudinal Feed Conversion Ratio. Lastly, we investigated whether FI, and therefore FCR, could be measured over a shorter period without compromising the description of the dynamic of FCR over time, to maintain the possibility to select for features of this dynamic. Reducing the time period for FI recording would allow collecting of records for more animals for this trait, and therefore FCR, and potentially increase the genetic gain. We first defined 3 different 5-wk periods with FCR records (initial period, wk 1 to 5; intermediate period, wk 3 to 7; and late period, wk 6 to 10) to be analyzed using the SAD model. The corresponding SBV were then computed as previously described for this model. The period providing the SBV with the highest correlation with the first and second SBV obtained from the genetic covariance matrix **G** with the SAD model (**SBV_SAD1** and **SBV_SAD2**) obtained over the entire 10-wk period was considered the best period for recording FCR.

Next, starting from the previous best 5-wk period, the number of weeks used in the analysis was increased by 1 wk at a time from 5 to 8 wk, and the same comparison was applied to determine the minimum number of weeks needed to provide a "satisfactory" SBV for FCR.

RESULTS

Estimation of Genetic Parameters

After selection, the RR model of degree 2 for genetic and permanent environmental effects was retained as the best model within the RR category and required 13 parameters to be estimated. Meanwhile, for the SAD approach, SAD1–22 was selected as the best SAD model for genetic and permanent environmental effects and required 12 parameters to be estimated. The BIC values for the best RR and SAD models were −9,838 and −10,060, respectively, indicating that the SAD approach provided the best fit for the data.

The changes in heritabilities over time are shown in Fig. 1. The heritability estimates were generally higher with the SAD model than with the RR model. They ranged from 0.22 to 0.46 (SE 0.03–0.06) for the SAD model and from 0.08 to 0.33 (SE 0.02–0.07) for the RR model. The heritabilities obtained with the RR model decreased up to wk 5 and then increased again toward the end of the test. For the SAD model, the heritability estimates were quite high at the beginning, decreased to a minimum at wk 8 (0.21 ± 0.03), and then increased before the end of the test period to reach values similar to the RR estimations. The ranges of SE were similar for the 2 approaches, from 0.02 to 0.07 for the RR model and from 0.03 to 0.06 for the SAD model.

The genetic correlations estimated for FCR over the 10 wk using the RR and SAD models are presented in Fig. 2. The genetic correlations between 2 given weeks depended on the time interval between the weeks. The shorter the interval, the higher the correlation. Correlations ranged from −0.39 to 0.98 for the RR model and from 0.08 to 0.83 for the SAD model. Consecutive week correlations were high and positive for both models, ranging from 0.91 to 0.98 for the RR

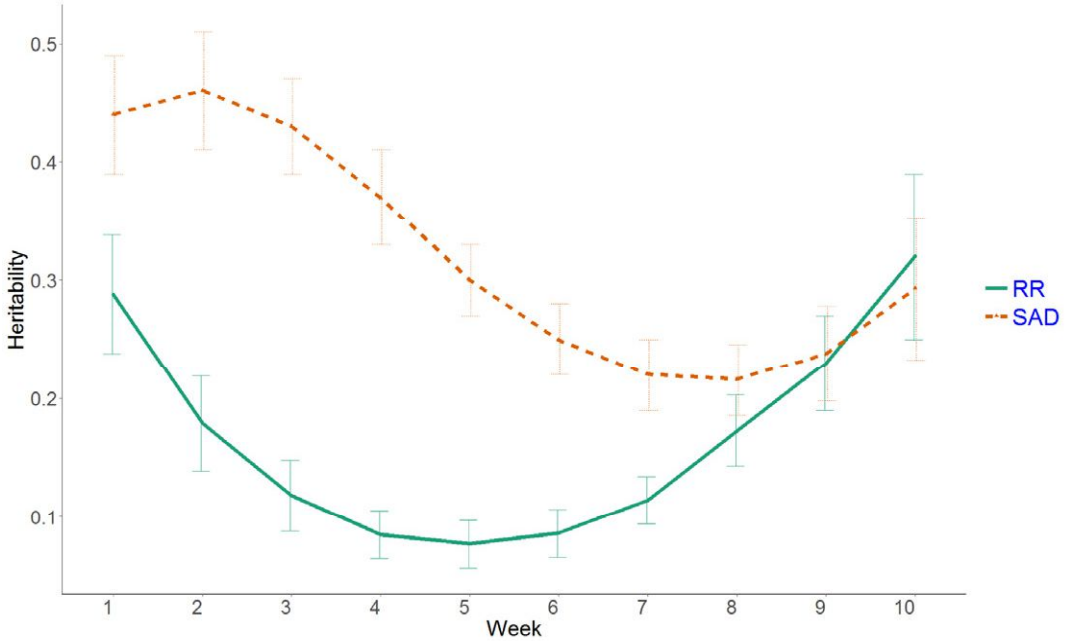

Figure 1. Changes of heritability estimates for feed conversion ratio over time under the random regression (RR) model using Legendre orthogonal polynomials and the structured antedependence (SAD) model. Standard errors are indicated as bars for each point estimate.

model and from 0.68 to 0.83 for the SAD model. For the RR model, the genetic correlations decreased as the interval between the weeks increased and became negative, resulting in negative correlations when the time interval between weeks was more than 5 to 6 wk. For the SAD model, the correlations decreased with the time interval but remained positive.

Selection Criterion

Estimated Breeding Value Trajectory Classification. The Spearman correlation between the weekly EBV_SAD and weekly EBV_RR over 10 wk for all animals was 0.95. The individual EBV trajectories under the 2 models (RR and SAD) were classified into 3 groups as shown in Fig. 3. The 3 patterns of EBV trajectories were similar for both models. Cohen's kappa agreement between the models was 0.80. The first EBV trajectory pattern was a continuous EBV increase over time with a weak slope and a low initial value (35.4 and 40.3% of the animals for the RR and SAD models, respectively; "A" group). The second pattern also reflected an increase of the EBV over time but had a steeper initial slope and higher initial value (34.6 and 32.6% of the animals for the RR and SAD models, respectively; "B" group). The last EBV trajectory pattern simply reflected a constant EBV over time (30.0 and 27.1% of the animals for the RR and SAD models, respectively; "C" group).

Selection Criterion Using Summarized Breeding Values. The approach based on eigendecomposition showed that the 2 first eigenvalues of $\mathbf{G}_{RR}$ and genetic covariance matrix for SAD model ($\mathbf{G}_{SAD}$) explained 90 and 73% of the genetic variation, respectively. The correlations between the SBV obtained with the different approaches are presented in Fig. 4. It should be noted that, depending on the program used to compute the eigendecomposition, matrices of eigenvectors of opposite signs can be obtained for the same initial correlation matrix. Therefore, we chose the signs of eigenvectors matrices to maximize the number of positive correlations with cEBV. The first summarized breeding values obtained from the matrix K with the RR model was highly correlated with SBV_RR1 (0.99) and SBV_SAD2(0.99), whereas the second SBV obtained from the coefficients covariance matrix **K** with the RR model (**SBV_RRK2**) was highly correlated with the second summarized breeding value obtained from the genetic covariance matrix **G** with the RR model (**SBV_RR2**; 0.99) and SBV_SAD1 (0.88) and also with sEBV_RR (0.96), sEBV_SAD (0.92), and cEBV (0.93).

In addition, the plots of the first 2 SBV depending on the trajectory clusters previously identified for the RR and SAD models (see Fig. 3) are presented in Fig. 5. For both approaches, the first 2 SBV were sufficient to describe the EBV trajectory types: for instance, for the SAD approach, animals in group A had low SBV_SAD1 values, animals in group B had high SBV_SAD1 and high SBV_SAD2 values, and animals in group C had high SBV_SAD1 and low SBV_SAD2 values. This suggested that SBV_SAD1 captured the average values of EBV over time, whereas SBV_SAD2 captured the slope of the EBV curve. The correspondences between SBV_RR*p* and EBV trajectories obtained with the RR

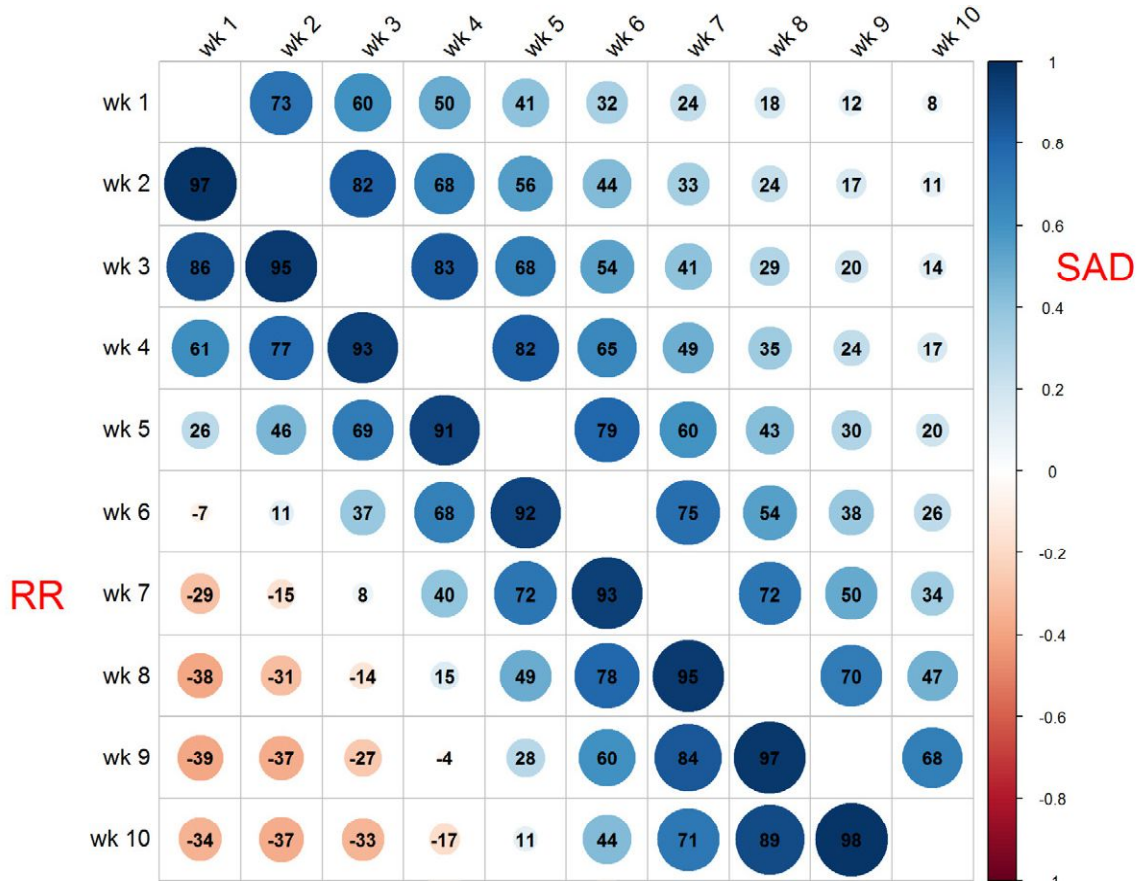

Figure 2. Genetic correlation estimates (x100) between times estimated with the random regression (RR) model (below the diagonal) and the structured antedependence (SAD) model (above the diagonal). The magnitude and sign of the correlations are indicated with darker and larger circles and blue (positive) or red (negative) colors, respectively.

approach also showed a clear distribution of the individuals from each group trajectory according to combinations of SBV_RR*p*. Finally, the plot of the eigenfunctions (Fig. 6) showed that the first eigenfunction was negative during the first 2 wk, then was positive from wk 2 until wk 9, and became negative again and decreased until the end of the test. The second eigenfunction was always positive and stable from wk 1 to 5, then increased, and reached a maximum at the end of the test.

Appropriate Period for Estimating Longitudinal Feed Conversion Ratio. The correlations between SBV_SAD1 and SBV_SAD2 obtained for reduced test periods and SBV_SAD1 and SBV_SAD2 obtained for the whole test period were estimated. The SBV_SAD1 related to the middle period had a higher correlation to SBV_SAD1 for the whole test period than that related to the first 5-wk period (0.93 vs. 0.89, respectively), whereas its correlation with the SBV_SAD2 was lower (0.67 vs. 0.69 for wk 1 to 5 and wk 3 to 7, respectively). When the evaluation period was extended by 1 wk toward the beginning or toward the end of the test period, these correlations did not increase for the middle period, contrary to those of the first 5-wk period extended for wk 6 (results not shown). Therefore, only results for the extended periods starting at the beginning of the test are reported. In this situation, the correlation between SBV_SAD1 (SBV_SAD2) for the reduced period and SBV_SAD1 (SBV_SAD2) for the whole test period increased with the number of weeks included, from 0.89 (0.69; wk 1 to 5) to 0.98 (0.87; wk 1 to 8; Fig. 7).

DISCUSSION

Estimation of Genetic Parameters

Using the BIC, the SAD model showed a slightly better fit to the data than the RR model. Furthermore, the predictive ability 1 wk ahead, computed as proposed by David et al. (2015), was similar for the 2 models (average Vonesh concordance coefficient = 0.39 for both). The SAD model provided higher heritability estimates than the RR model. Similar results have been found in the literature for other traits (Jaffrézic et al., 2004; David et al., 2015). The lower values of heritability obtained with the RR model might be a consequence of the border effect problem associated with this model, which is eliminated in

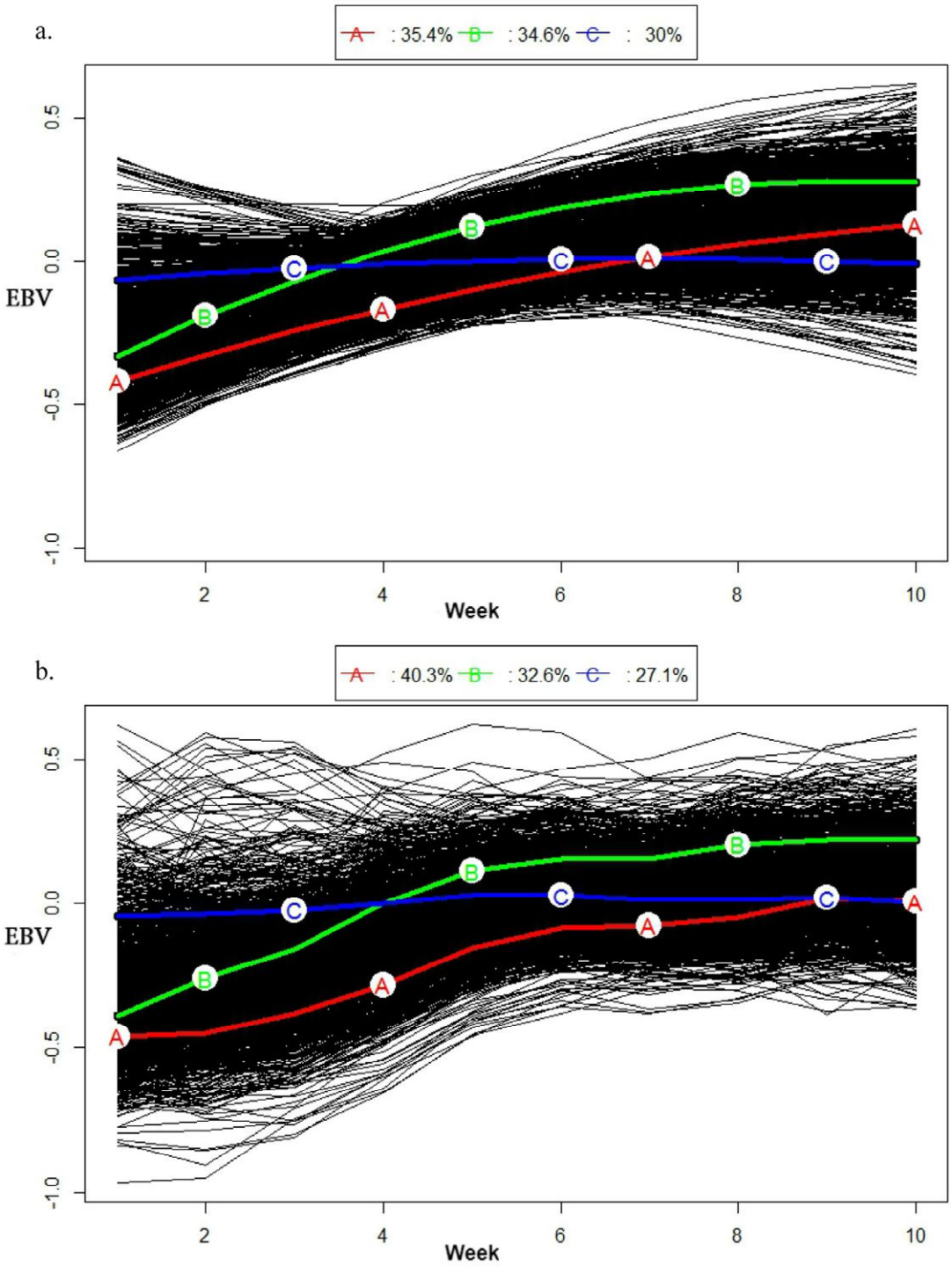

Figure 3. Individual EBV trajectories (in black) and group trajectories resulting from nonhierarchical *k*-means clustering analyses with 3 groups obtained with the random regression (a) and structured antedependence models (b). The proportion of individuals gathered in each group is indicated above each graph.

the SAD model that combines the antedependence parameters and innovation variances (Jaffrézic et al., 2004), suggesting a greater confidence in the genetic parameters obtained with the SAD model. This was reinforced by results from a multiple trait model with a diagonal covariance matrix applied to weekly FCR. Heritabilities obtained with the SAD model were closer to those of the multiple trait model than heritabilities of the RR model with this multiple trait model (average absolute difference = 0.09 vs. 0.15, respectively), the heritabilities being systematically lower with the RR model. Nonetheless, it should be noted that the computing time of the SAD model for each iteration is longer than the one of the RR model (2.7 times longer, on average) but SAD models generally converge with fewer iterations. Consequently, on our data set, the total computing time of the SAD model was 1.2 times longer than for the RR model.

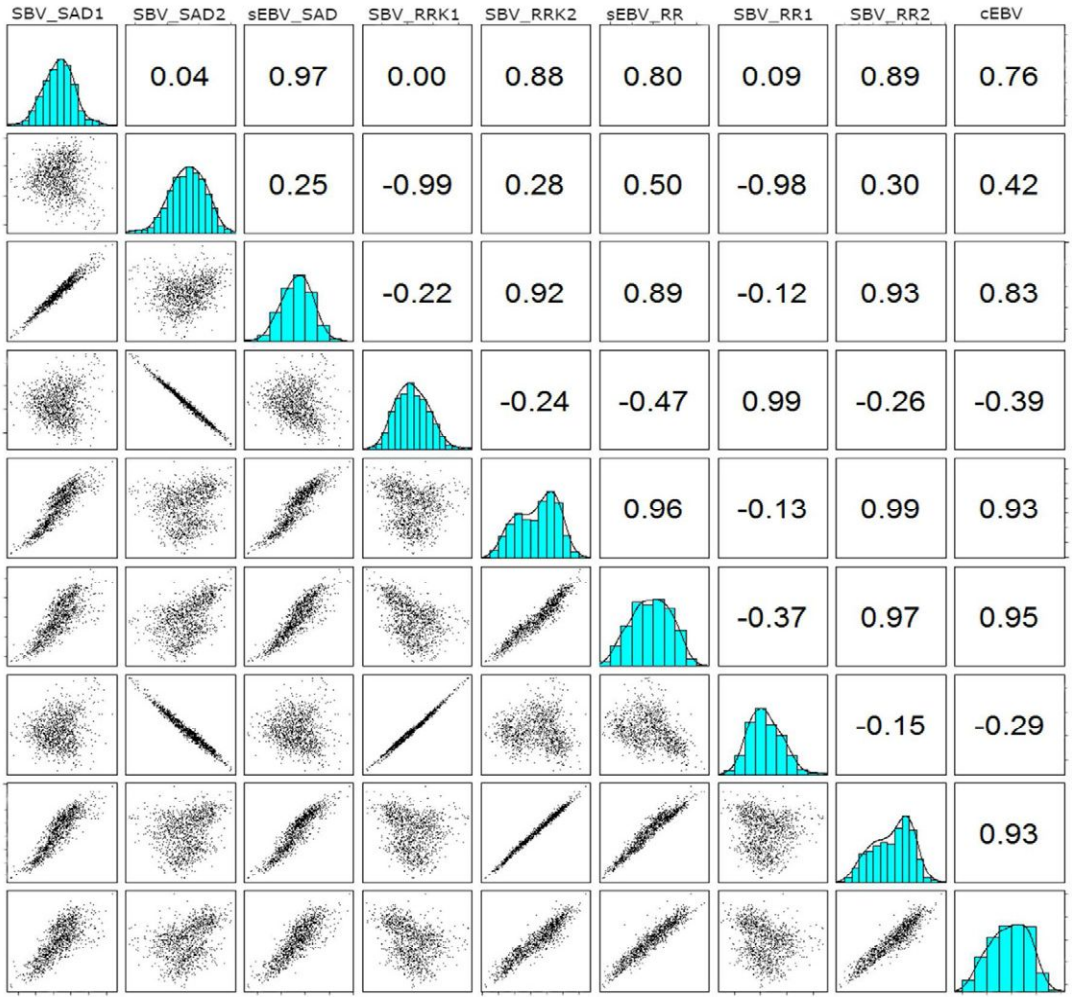

Figure 4. Distributions of the summarized breeding values (SBV_SAD1 and SBV_SAD2 = first and second summarized breeding values, respectively, obtained from the genetic covariance matrix **G** with the structured antedependence [SAD] model; SBV_RRK1 and SBV_RRK2 = first and second summarized breeding values, respectively, obtained from the coefficients covariance matrix **K** with the random regression [RR] model; SBV_RR1 and SBV_RR2 = first and second summarized breeding values, respectively, obtained from the genetic covariance matrix **G** with the RR model), the sums of EBV over the 10 wk (sEBV_RR and sEBV_SAD = sum of EBV obtained with the RR and SAD models, respectively), and the EBV from the animal model using the feed conversion ratio computed over the 10-wk period (cEBV; on the diagonal), joint distributions of these estimates (below the diagonal), and Spearman correlations between the estimates (above the diagonal).

We used a bootstrap procedure to compute the SE of heritability obtained with the SAD model. It is also feasible to use Taylor expansion to obtain an approximate SE. Nonetheless, the formula becomes complex when the order of the antedependence increases.

The heritabilities obtained with the SAD model (from 0.22 to 0.46) at different weeks were in line with those reported in the literature for FCR values recorded over the full growing period on earlier generations of the same population (0.24 ± 0.06; Saintilan et al., 2012) and in other Large White/Yorkshire populations: 0.26 ± 0.07 (Bunter et al., 2010), 0.30 ± 0.03 (Saintilan et al., 2013), and 0.32 ± 0.05 (Do et al., 2013). The changes of heritabilities with time are consistent with the assumption that different genes can be associated with FCR at different stages of growth, as suggested by Shirali et al. (2013) for residual FI and FCR.

For the SAD and RR models, the genetic correlations decreased as the time interval between measurements increased. They became negative in the case of the RR model, although this is unlikely to reflect the true correlations between these distant periods. It has been previously reported that because the RR model cannot handle correlations that asymptotically tend to 0, it provides biased estimates of the correlations for distant time intervals (Jaffrézic et al., 2004).

In such cases, the correlations become negative, as observed in previous studies (David et al., 2015). It should be noted that considering heterogeneous residual variance with time in the RR model did not modify these negative value estimates and did not reduce the border effects problem (results not shown). The positive genetic correlations over time estimated with the SAD model suggest that efficient animals with low FCR values at the beginning of the test period tend

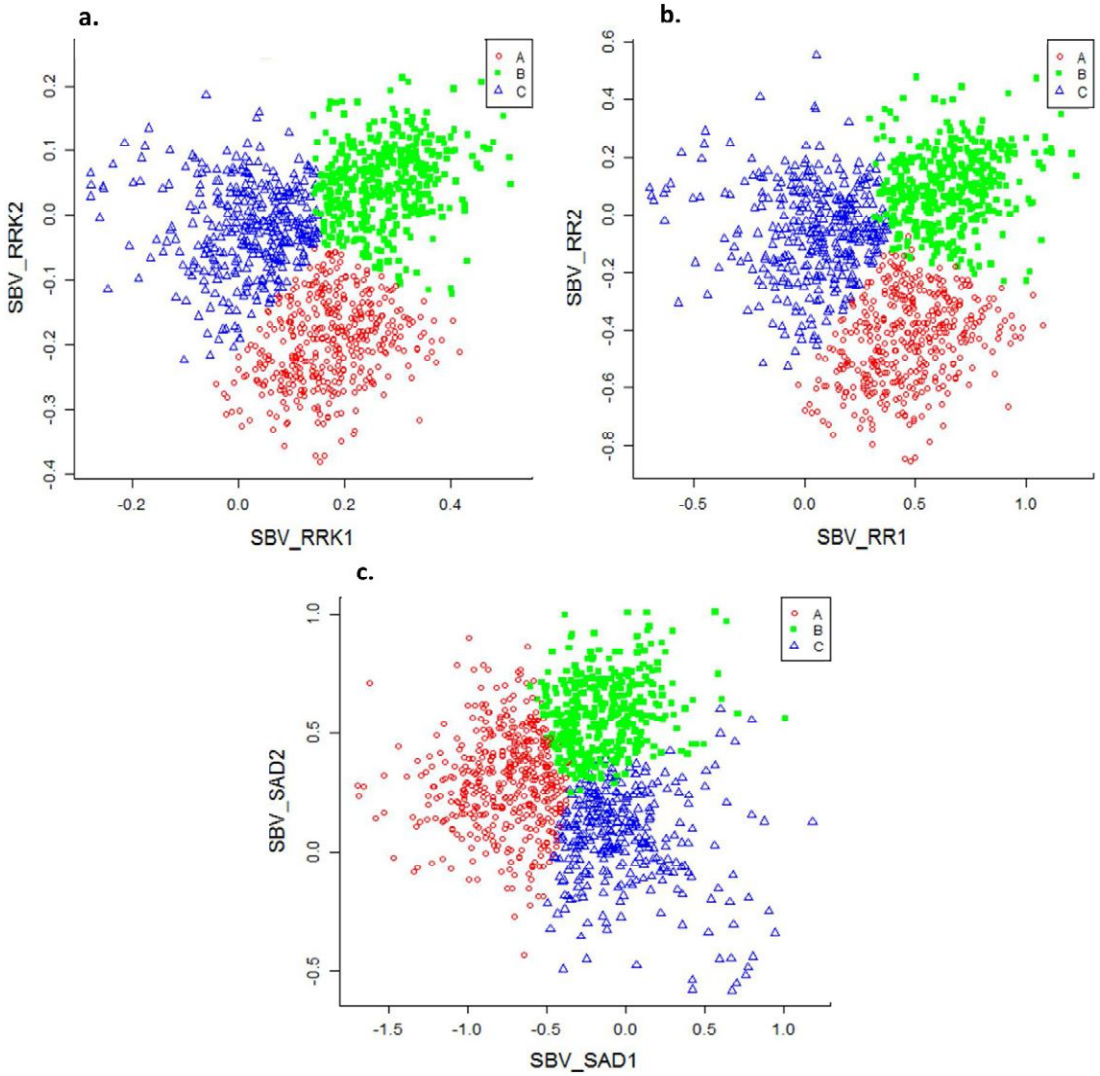

Figure 5. Scatterplots of the individual first and second summarized breeding values obtained from the estimated covariance matrices **K** with the random regression (RR) model (a; SBV_RRK1 and SBV_RRK2, respectively), from the genetic covariance matrix **G** with the RR model (b; SBV_RR1 and SBV_RR2, respectively), and from the genetic covariance matrix **G** with the structured antedependence (SAD) model (c; SBV_SAD1 and SBV_SAD2, respectively). The groups of trajectories to which each individual belongs as determined using the nonhierarchical *k*-means approach (see Fig. 3) applied to the longitudinal EBV from the RR model (a and b) and from the SAD model (c) are indicated as red circles (A group), green squares (B group), and blue triangles (C group).

to also have a lower FCR toward the middle of the test, but more independent results seem to be expected toward the end of the test. Henryon et al. (2002) estimated the genetic correlations between FCR values in growing rainbow trout at different time points over a 215-d period. Most of the correlations were positive and high. No reports about such genetic correlations for FCR could be found in the literature for pigs.

Selection Criterion

Modeling longitudinal data yields more accurate EBV due to the inclusion of repeated records over time and consideration of the covariance structure of the data (Boligon et al., 2011). However, the main difficulty of selection based on repeated measurement analysis is the obtention of as many EBV as time points used for the evaluation. The general idea is, therefore, to summarize these multiple EBV into a smaller set of new composite dimensions with a minimum loss of information (Van Der Werf et al., 1998), as successfully applied by Buzanskas et al. (2013). Ideally, 1 or 2 indexes can capture the individual EBV trajectory profiles to ease animal selection.

In the current study, a classification approach was used to identify different typical EBV trajectories from the SAD and RR approaches, as earlier proposed to cluster egg production curves at the phenotypic level by Savegnago et al. (2011) and milk yield profiles at the genetic level by Savegnago et al. (2016). This trajectory classification is proposed in our study as a complementary analysis to describe the group trajectories and better comprehend the animal profiles as compared with the selection objectives of a breeding

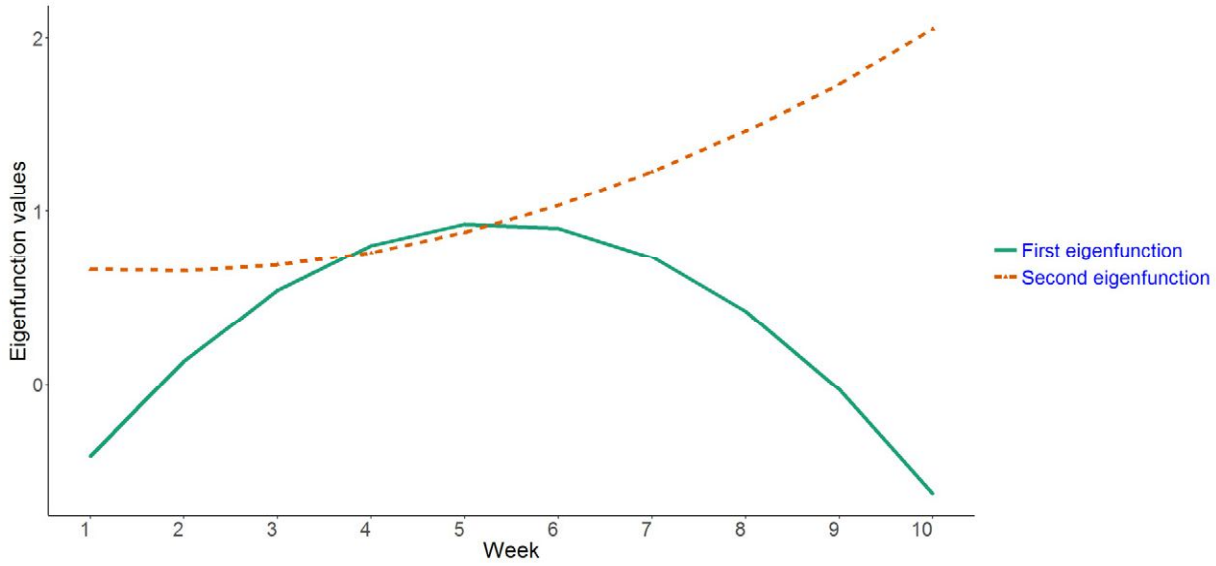

Figure 6. Two first eigenfunctions (unitless) associated with the covariance matrix **K** of the random regression model represented over the 10 wk of the test.

program. To summarize the EBV, we applied an eigendecomposition of the **G** genetic covariance matrices from the RR and SAD models, as originally applied to the **K** matrix of the RR models (Van Der Werf et al., 1998). The eigendecomposition has the advantage of accounting for the genetic covariances between weeks, which is not the case when using the average of the weekly EBV. By extracting the main axes of covariability among the EBV along time, 1 or 2 eigencomponents usually capture almost all the additive genetic variation in level and shape of the genetic curve, at least when applied to lactation curves (Druet et al., 2005; Togashi and Lin, 2006). Our results show that the first 2 SBV obtained from the eigendecomposition of the **G** matrix of the SAD model provided information similar to that of the eigendecomposition of the **K** matrix from the RR model. These SBV could, therefore, be similarly interpreted based on the eigenfunctions from the **K** matrix or the trajectory classification applied to the weekly EBV. As a result, combinations of the 2 first SBV are sufficient to describe the 3 groups of trajectories. It suggests that animals within a trajectory share genetic features that drive the dynamics of their feed efficiency during growth.

Despite differences in the estimation of the genetic parameters between the 2 approaches, the selection results were very similar for the RR and SAD models and could be confirmed by computing correlations between different SBV and with cEBV. In practice, one of the SBV was related to the average level of FCR during the test period (SBV_SAD1 and SBV_RR2) and the other one was related to the slope of the curve over time (SBV_RR1 and SBV_SAD2). In spite of this high concordance between the 2 approaches, the first 2 eigenvectors according to $\mathbf{G}_{SAD}$ explained only about 73% of the genetic variation, which is rather lower than for the RR model (90%).

As expected from earlier studies (Kirkpatrick et al., 1990; Meyer and Kirkpatrick, 2005), the use of the **K** matrix and the **G** matrix of the RR model to calculate SBV led to very similar results. In the present study, the first eigenfunction changed sign with time. This suggests that selection for this first component would have opposite effects for the intermediate period compared with the extreme periods (2 wk at the beginning and 2 wk at the end of the trajectory). The second eigenfunction increased with time and was always positive. This means that selection for SBV_RRK2 would lead to selection in the same direction for all the time points, with a higher weight at the end of the testing period in comparison with the beginning of the period. Due to the high correlation between the first SBV obtained from the coefficients covariance matrix **K (SBV_RRK1)** and SBV_RR1 or SBV_SAD2, it confirmed our interpretation of SBV_RR1 and SBV_SAD2 as indicators of the slope of the feed efficiency curve.

To summarize, SBV can be used for selection purposes. To fully evaluate their potential, the estimation of genetic correlations with other production traits would provide a better insight on the use of the trajectories for selection. Indeed, it can be assumed that animals from the A group (low average FCR but a regular increase over time) would show a different fat content at slaughter than animals from the C group of similar average FCR, so selection for different FCR trajectories would consolidate breeding objectives on carcass composition. Further comparison of responses to selection for the traits of the breeding objective using different indexes options (cEBV and two first SBVs associated to the two first eigenvector of the matrix G)

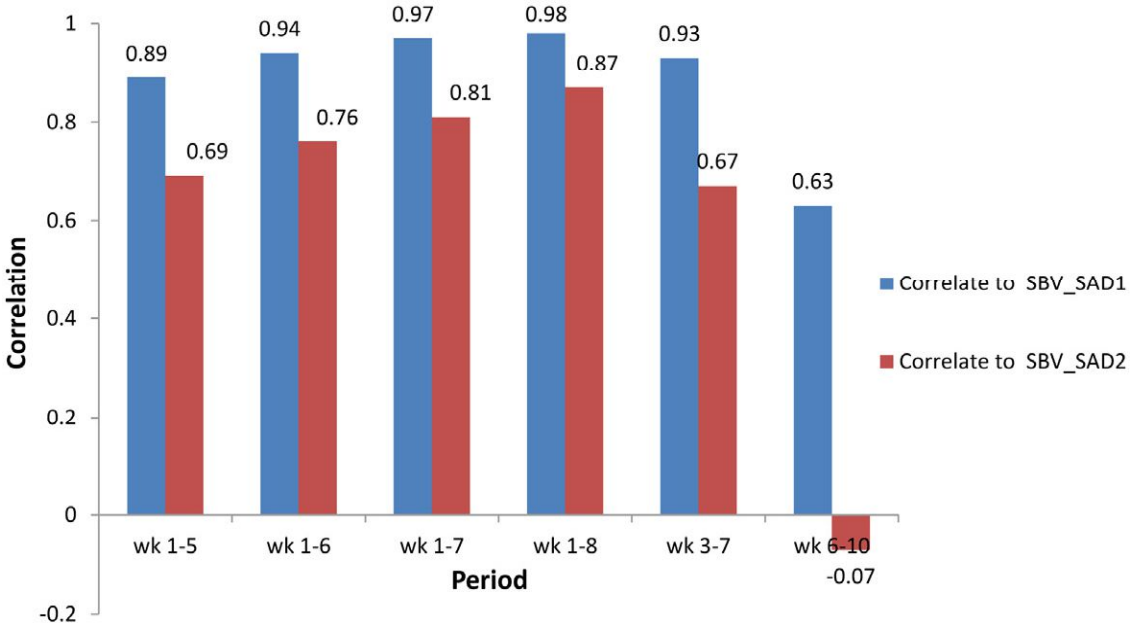

Figure 7. Correlations between the first (blue bars) and second (red bars) summarized breeding values obtained from the genetic covariance matrix **G** with the structured antedependence model (SBV_SAD1 and SBV_SAD2, respectively) for the full test period and those obtained from records from wk 1 to 5, wk 1 to 6, wk 1 to 7, wk 1 to 8, wk 3 to 7, and wk 6 to 10.

or a combination of two among them, would clarify the possible selection strategies.

Appropriate Period for Estimating Longitudinal Feed Conversion Ratio

The accuracy of the estimation of genetic parameters heavily relies on the quantity of data available. On the other hand, the cost of individual FI measures is high. Therefore, there is a trade-off between parsimony, complexity of the analysis, and potential bias, so choices need to be made. The goal is, therefore, to reduce the duration of the test period for FI with a minimum loss of accuracy for animal selection for FCR dynamic features, to test more pigs and increase the genetic gain (Begli et al., 2016). Wetten et al. (2012) proposed to use information on early periods of FI combined with information on growth to reduce the test period. In the current study, a similar conclusion was reached: the first weeks of test showed better correlations to selection criteria obtained with the whole test period than the middle and the last periods. The selection accuracy could be increased stepwise by extending the evaluation from 5 to 8 wk of duration. Further studies are required to better understand the link between the genetic gain, the costs associated with different strategies, and the changes in prediction accuracy due to a combined reduction of the duration of the test period and a greater number of pigs tested.

Conclusion

The current study indicates that the SAD model is promising for genetic selection: 1) it requires fewer parameters to fit the covariance matrices than the RR model and 2) it is not associated with the border effect problems and negative correlation estimates observed with the RR model. The use of SBV is a solution for animal selection applicable with the SAD model. The results of this study also suggest that a reduction of the duration of the FI test period to reduce measurement costs is probably feasible to select for feed efficiency.

LITERATURE CITED

Begli, H. E., R. V. Torshizi, A. A. Masoudi, A. Ehsani, and J. Jensen. 2016. Longitudinal analysis of body weight, feed intake and residual feed intake in F_2 chickens. Livest. Sci. 184:28–34. doi:10.1016/j.livsci.2015.11.018

Boligon, A. A., F. Baldi, M. E. Z. Mercadante, R. B. Lobo, R. J. Pereira, and L. G. Albuquerque. 2011. Breeding value accuracy estimates for growth traits using random regression and multi-trait models in Nelore cattle. Genet. Mol. Res. 10:1227–1236. doi:10.4238/vol10-2gmr1087

Bunter, K. L., W. Cai, D. J. Johnston, and J. C. M. Dekkers. 2010. Selection to reduce residual feed intake in pigs produces a correlated response in juvenile insulin-like growth factor-I concentration. J. Anim. Sci. 88:1973–1981. doi:10.2527/jas.2009-2445

Buzanskas, M. E., R. P. Savegnago, D. A. Grossi, G. C. Venturini, S. A. Queiroz, L. O. C. Silva, R. A. A. Torres Júnior, D. P. Munari, and M. M. Alencar. 2013. Genetic parameter estimates and principal component analysis of breeding values of reproduction and growth traits in female Canchim cattle. Reprod. Fertil. Dev. 25:775–781. doi:10.1071/RD12132

David, I., H. Garreau, E. Balmisse, Y. Billon, and L. Canario. 2017. Multiple-trait structured antedependence model to study the relationship between litter size and birth weight in pigs and rabbits. Genet. Sel. Evol. 49:11. doi:10.1186/s12711-017-0288-3

David, I., J. Ruesche, L. Drouilhet, H. Garreau, and H. Gilbert. 2015. Genetic modeling of feed intake. J. Anim. Sci. 93:965–977. doi:10.2527/jas.2014-8507

Do, D. N., A. B. Strathe, J. Jensen, T. Mark, and H. N. Kadarmideen. 2013. Genetic parameters for different measures of feed efficiency and related traits in boars of three pig breeds. J. Anim. Sci. 91:4069–4079. doi:10.2527/jas.2012-6197

Druet, T., F. Jaffrézic, and V. Ducrocq. 2005. Estimation of genetic parameters for test day records of dairy traits in the first three lactations. Genet. Sel. Evol. 37:257–271. doi:10.1186/1297-9686-37-4-257

Efron, B., and T. Hastie. 2016. The jackknife and bootstrap. In: Computer age statistical inference: Algorithms, evidence, and data science. Cambridge Univ. Press, Cambridge, UK. p. 155–180. doi:10.1017/CBO9781316576533.011

Englishby, T. M., G. Banos, K. L. Moore, M. P. Coffey, R. D. Evans, and D. P. Berry. 2016. Genetic analysis of carcass traits in beef cattle using random regression models. J. Anim. Sci. 94:1354–1364. doi:10.2527/jas.2015-0246

Fischer, T. M., A. R. Gilmour, and J. H. J. van der Werf. 2004. Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. Genet. Sel. Evol. 36:363–369. doi:10.1186/1297-9686-36-3-363

Genolini, C., X. Alacoque, M. Sentenac, and C. Arnaud. 2015. kml and kml3d: R packages to cluster longitudinal data. J. Stat. Softw. 65:1–34. doi:10.18637/jss.v065.i04

Gilbert, H., J.-P. Bidanel, J. Gruand, J.-C. Caritez, Y. Billon, P. Guillouet, H. Lagant, J. Noblet, and P. Sellier. 2007. Genetic parameters for residual feed intake in growing pigs, with emphasis on genetic relationships with carcass and meat quality traits. J. Anim. Sci. 85:3182–3188. doi:10.2527/jas.2006-590

Gilbert, H., Y. Billon, L. Brossard, J. Faure, P. Gatellier, F. Gondret, E. Labussière, B. Lebret, L. Lefaucheur, N. L. Floch, I. Louveau, E. Merlot, M.-C. Meunier-Salaün, L. Montagne, P. Mormede, D. Renaudeau, J. Riquet, C. Rogel-Gaillard, J. van Milgen, A. Vincent, and J. Noblet. 2017. Review: Divergent selection for residual feed intake in the growing pig. Animal 11(9):1427–1439. doi:10.1017/S175173111600286X

Gilmour, A. R., B. J. Gogel, and B. R. Cullis. 2009. ASReml user guide release 3.0. VSN Int. Ltd., Hemel Hempstead, UK.

Henryon, M., A. Jokumsen, P. Berg, I. Lund, P. B. Pedersen, N. J. Olesen, and W. J. Slierendrecht. 2002. Genetic variation for growth rate, feed conversion efficiency, and disease resistance exists within a farmed population of rainbow trout. Aquaculture 209:59–76. doi:10.1016/S0044-8486(01)00729-3

Huynh-Tran, V. H., H. Gilbert, and I. David. 2017. How to improve breeding value prediction for feed conversion ratio in the case of incomplete longitudinal body weights. J. Anim. Sci. 95:39–48. doi:10.2527/jas.2016.0980

Jaffrézic, F., and S. D. Pletcher. 2000. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. Genetics 156:913–922.

Jaffrézic, F., E. Venot, D. Laloë, A. Vinet, and G. Renand. 2004. Use of structured antedependence models for the genetic analysis of growth curves. J. Anim. Sci. 82:3465–3473. doi:10.2527/2004.82123465x

Kirkpatrick, M., D. Lofsvold, and M. Bulmer. 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics 124:979–993.

Meyer, K. 2005. Random regression analyses using B-splines to model growth of Australian Angus cattle. Genet. Sel. Evol. 37:473. doi:10.1186/1297-9686-37-6-473

Meyer, K., and W. G. Hill. 1997. Estimation of genetic and phenotypic covariance functions for longitudinal or "repeated" records by restricted maximum likelihood. Livest. Prod. Sci. 47:185–200. doi:10.1016/S0301-6226(96)01414-5

Meyer, K., and M. Kirkpatrick. 2005. Up hill, down dale: Quantitative genetics of curvaceous traits. Philos. Trans. R. Soc., B 360:1443–1455. doi:10.1098/rstb.2005.1681

Patience, J. F., M. C. Rossoni-Serão, and N. A. Gutiérrez. 2015. A review of feed efficiency in swine: Biology and application. J. Anim. Sci. Biotechnol. 6:33. doi:10.1186/s40104-015-0031-2

Saintilan, R., I. Mérour, L. Brossard, T. Tribout, J. Y. Dourmad, P. Sellier, J. Bidanel, J. van Milgen, and H. Gilbert. 2013. Genetics of residual feed intake in growing pigs: Relationships with production traits, and nitrogen and phosphorus excretion traits. J. Anim. Sci. 91:2542–2554. doi:10.2527/jas.2012-5687

Saintilan, R., P. Sellier, Y. Billon, and H. Gilbert. 2012. Genetic correlations between males, females and castrates for residual feed intake, feed conversion ratio, growth rate and carcass composition traits in Large White growing pigs. J. Anim. Breed. Genet. 129:103–106. doi:10.1111/j.1439-0388.2011.00972.x

Savegnago, R. P., S. L. Caetano, S. B. Ramos, G. B. Nascimento, G. S. Schmidt, M. C. Ledur, and D. P. Munari. 2011. Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population. Poult. Sci. 90:2174–2188. doi:10.3382/ps.2011-01474

Savegnago, R. P., G. B. do Nascimento, G. J. de Magalhães Rosa, R. L. R. de Carneiro, R. C. Sesana, L. El Faro, and D. P. Munari. 2016. Cluster analyses to explore the genetic curve pattern for milk yield of Holstein. Livest. Sci. 183:28–32. doi:10.1016/j.livsci.2015.11.010

Schnyder, U., A. Hofer, F. Labroue, and N. Künzi. 2001. Genetic parameters of a random regression model for daily feed intake of performance tested French Landrace and Large White growing pigs. Genet. Sel. Evol. 33:635–658. doi:10.1186/1297-9686-33-6-635

Schwarz, G. 1978. Estimating the dimension of a model. Ann. Stat. 6:461–464. doi:10.1214/aos/1176344136

Shirali, M., C.-A. Duthie, A. Doeschl-Wilson, P. W. Knap, E. Kanis, J. A. van Arendonk, and R. Roehe. 2013. Novel insight into the genomic architecture of feed and nitrogen efficiency measured by residual energy intake and nitrogen excretion in growing pigs. BMC Genet. 14:121. doi:10.1186/1471-2156-14-121

Togashi, K., and C. Y. Lin. 2006. Selection for milk production and persistency using eigenvectors of the random regression coefficient matrix. J. Dairy Sci. 89:4866–4873. doi:10.3168/jds.S0022-0302(06)72535-8

Van Der Werf, J. H., M. E. Goddard, and K. Meyer. 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. J. Dairy Sci. 81:3300–3308. doi:10.3168/jds.S0022-0302(98)75895-3

Wang, W. 2013. Identifiability of linear mixed effects models. Electron. J. Stat. 7:244–263. doi:10.1214/13-EJS770

Wetten, M., J. Ødegård, O. Vangen, and T. H. E. Meuwissen. 2012. Simultaneous estimation of daily weight and feed intake curves for growing pigs by random regression. Animal 6:433–439. doi:10.1017/S1751731111001832