Archivos de Zootecnia
Journal website: https://www.uco.es/ucopress/az/index.php/az/
SHORT NOTE

Genetic analysis of pig growth curves: a simulation study

Marques, D.B.D.@; Silva, F.F.; Veroneze. R.; Lopes, P.S. and Guimarães, S.E.F.

Departamento de Zootecnia. Universidade Federal de Viçosa. Viçosa. Minas Gerais. Brazil.

SUMMARY

We aimed to estimate genetic parameters for the non-linear Gompertz growth curve parameters based on the two-step method and simulated data of a pig population. Data of true breeding values, growth curve parameters (a, b and k) and body weights at 1, 90, 210, 360 and 540 days of age were simulated for 1700 individuals. Afterwards, in the first stage of the two-step method, the Gompertz curves were fitted for each animal and the growth curve parameters were estimated. In the second step, considering the estimated parameters as phenotypes, animals' breeding values were predicted by multitrait analyses. The simulation and two-step processes were replicated 10 times. For each replicate, the genetic parameters, accuracies and biases were calculated. Moderate to high average heritabilities for the curve parameters [0.62 (0.04), 0.56 (0.04) and 0.57 (0.04) for a, b and k, respectively] show the possibility of effective selection of animals based on these parameters. High and favorable average genetic correlations, especially between the most important growth curve parameters [a and k, 0.93 (0.01)], shows that selection for increased mature weight (a) will also increase maturation rate (k). High average accuracies [0.84 (0.02), 0.82 (0.02) and 0.83 (0.02)] and low average biases [0.06 (0.06), 0.06 (0.06) and 0.07 (0.07) for a, b, and k, respectively] indicate that the two-step approach is feasible for predicting breeding values for growth curve parameters and for dealing with longitudinal body weight measurements.

ADDITIONAL KEYWORDS
Accuracy.
Bias.
Genetic parameters.
Gompertz.
Non-linear models..

Análise genética de curvas de crescimento de suínos: um estudo de simulação

RESUMO

Objetivamos estimar parâmetros genéticos para os parâmetros da curva de crescimento não-linear de Gompertz com base no método em duas etapas e em dados simulados de uma população de suínos. Dados de valores genéticos reais, dos parâmetros da curva de crescimento (a, b e k) e dos pesos corporais aos 1, 90, 210, 360 e 540 dias de idade foram simulados para 1700 indivíduos. Posteriormente, na primeira etapa do método em duas etapas, as curvas de Gompertz foram ajustadas para cada animal e os parâmetros da curva de crescimento foram estimados. Na segunda etapa, considerando os parâmetros estimados como fenótipos, os valores genéticos dos animais foram preditos por análises multicaracterísticas. Os processos de simulação e de duas etapas foram replicados 10 vezes. Para cada réplica, os parâmetros genéticos, acurácias e vieses foram calculados. Herdabilidades médias moderadas a altas para os parâmetros da curva [0,62 (0,04), 0,56 (0,04) e 0,57 (0,04) para a, b e k, respectivamente] mostram a possibilidade de seleção eficaz de animais com base nesses parâmetros. Correlações genéticas médias altas e favoráveis, especialmente entre os parâmetros mais importantes da curva de crescimento [a e k, 0,93 (0,01)], mostram que a seleção para o aumento do peso à maturidade (a) também aumentará a taxa de maturação (k). Acurácias médias altas [0,84 (0,02), 0,82 (0,02) e 0,83 (0,02)] e vieses médios baixos [0,06 (0,06), 0,06 (0,06) e 0,07 (0,07) para a, b e k, respectivamente] indicam que a abordagem em duas etapas é viável para predizer valores genéticos para os parâmetros da curva de crescimento e para trabalhar com medidas longitudinais de peso corporal.

PALAVRAS CHAVE ADICIONAIS
Acurácia.
Gompertz.
Modelos não-lineares.
Parâmetros genéticos.
Viés.

INFORMATION
Cronología del artículo.
Recibido/Received: 07.07.2020.
Aceptado/Accepted: 07.11.2020
On-line: 15/04/2021
Correspondencia a los autores/Contact e-mail:
danielebdiniz@gmail.com

INTRODUCTION

The profitability and success of pig production systems depend on the knowledge of animals' body weight behavior over time, which can be achieved by analyzing the animals' growth curves (Silva et al., 2013). This knowledge allows evaluating the animals' growth profile in the current production system (Silva et al., 2013) and the possibility of using alternative management systems, besides directing the decisions that must be made, mainly related to nutritional management and selection programs for growth traits.

According to Freitas (2005), growth curves in animal production relate the weights and ages of animals using, for example, non-linear models, and can be applied to i) summarize the growth traits in a few parameters of the non-linear model; ii) evaluate the profile of responses to treatments over time and iii) identify the best animals in a population. In general, non-linear models are fitted to predict the dependent variable (weight) for different values of the independent variable (ages). Besides the growth curve parameters that are directly obtained and present biological interpretation (such as mature weight and maturation rate), their genetic parameters (heritabilities and genetic correlations) can be estimated and considered in animal breeding programs (Silva, 2008). Therefore, using the estimates of the non-linear models parameters as phenotypes in genetic models, the animals breeding values (EBVs) for these parameters can be predicted and the breeding goals are achieved by changing the shape of the growth curves (Lázaro et al., 2017).

In pigs, the Gompertz non-linear function (Gompertz, 1825) has been used to accurately describe the potential growth of animals according to age (Wellock et al., 2004; Freitas, 2005; Koivula et al., 2008; Cai et al., 2012; Coyne et al., 2017; Lázaro et al., 2017). Gompertz function requires the determination of three parameters: mature body weight (a), the inflection point (b) and the maturation rate (k). With these parameters, it is possible to understand the biological growth-related process across the growth trajectory (Cai et al., 2012).

In this context, we aimed to estimate the genetic parameters for the non-linear Gompertz growth curve parameters based on the traditional two-step method and simulated data of a pig population.

MATERIAL AND METHODS

Data simulation was performed in R software (R Core Team, 2013), considering the following multitrait (3x3) linear model:

$$\mathbf{w} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \qquad \text{(I)}$$

in which: **w** is the vector of simulated Gompertz curve parameters (a, b and k); **β** is the vector of fixed effects (general mean); **u** is the vector of random true breeding values (TBV); **e** is the vector of random errors; and **X** and **Z** are the incidence matrices of **β** and **u**, respectively. It was assumed that $\mathbf{u} \sim N(\mathbf{0}, \mathbf{A} \otimes \Sigma_u)$ and $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I} \otimes \Sigma_e)$, in which ⊗ is a direct product operation; $\Sigma\mathbf{u}$ and Σ_e represent (co) variance matrices for **u** and **e**, respectively; **A** is the average numerator relationship matrix; and **I** is an identity matrix.

The pedigree was simulated using the *simul.pedigree* function of synbreed package (Wimmer et al., 2012), considering 3 generations and 100 individuals per generation. The number of individuals in each full-sib family in the last generation was assumed as 15, resulting in a total of 1700 animals. The **A** matrix was obtained based on the simulated pedigree using the *makeA* function of pedigree package (Coster, 2012). Afterwards, the **u** and **e** vectors were generated by the *mvrnorm* function of MASS package (Venables & Ripley, 2002), considering the genetic and residual (co) variance matrices (Σ_u and Σ_e, respectively) from Cai et al. (2012) study, as follows:

Similarly, the **β** vector was generated based on a, b and k mean values from Cai et al. (2012). Finally,

$$\Sigma\mathbf{u} = \begin{bmatrix} 7872.00 & 3499.05 & 2609.79 \\ 3499.05 & 1653.00 & 1221.36 \\ 2609.79 & 1221.36 & 979.20 \end{bmatrix}$$

$$\Sigma\mathbf{e} = \begin{bmatrix} 4428.00 & 2184.84 & 1556.94 \\ 2184.84 & 1197.00 & 849.04 \\ 1556.94 & 849.04 & 652.80 \end{bmatrix}$$

based on the simulated parameters of model (I), the growth curve parameters for each animal (a_i, b_i and k_i) were generated.

The next step in simulation process was to generate the body weights at 1, 90, 210, 360 and 540 days of age (Cai et al., 2012) for each pig. For that, the following Gompertz non-linear model was applied, considering the growth curve parameters (a_i, b_i and k_i) simulated in the previous step:

$$y_{ij} = a_i \ \exp\{-\exp[-(t_j - b_i) / k_i]\} + \varepsilon_{ij},$$

in which: y_{ij} is the vector of simulated body weights for animal i (i = 1, 2, …, n) measured on age t_j (j = 1,2,…); (j=1,2,...ri); ε_{ij} is the residual term, assumed as $\varepsilon ij \sim N(0, \sigma^2\varepsilon)$ and a_i, b_i and k_i are as described before.

In this context, n is the total number of animals; r_i is the is the number of repeated measures reported for animal i; and and $\sigma^2\varepsilon$ is the residual variance (assumed as 3.92, according to Cai et al., 2012).

The two-step method was applied in the current study. In this analysis, growth curve parameters are estimated based on simulated body weight data of each animal in the first stage; and, in the second stage, these parameters estimates are used in the genetic evaluation. Therefore, considering the simulated body weights at 1, 90, 210, 360 and 540 days of age, the Gompertz curves were fitted for each animal using the *nlsList* function of nlme package (Pinheiro et al., 2015) in the first stage. The estimated growth curve parameters for each animal i ($\hat{a}_i$, $\hat{b}_i$ and $\hat{k}_i$) were obtained from this first step and posteriorly used as dependent variables under a mixed model framework. In this sense, in the second step, EBVs were predicted and genetic parameters (heritabilities and genetic correlations) were estimated through multitrait analyses (3x3) in ASReml 3.0 software (Gilmour et al., 2009), considering $\hat{a}_i$, $\hat{b}_i$ and $\hat{k}_i$ as observed phenotypes. The simulation and the respective two-step process proposed here were replicated 10 times.

The simulated TBVs and EBVs were used to estimate the prediction accuracy and bias in each replicate. Accuracy was calculated as the Pearson correlation coefficient between TBVs and EBVs. The regression coefficients (β_1 from linear regression of TBV over EBV were used to calculate the bias, which was defined as the deviation from β_1 in terms of unity. Averages and standard deviations (SD) of accuracies, biases and genetic parameters were calculated considering the 10 replicates.

RESULTS

Average heritability estimates for Gompertz growth curve parameters ranged from 0.56 (0.04) for b estimated parameter to 0.62 (0.04) for a estimated parameter. Average genetic correlations were of high magnitudes, ranging from 0.93 (0.01) to 0.97 (0.01) between a and k, and between a and b estimated parameters, respectively (**Table I**).

Average accuracies were of high magnitudes, ranging from 0.82 (0.02) to 0.84 (0.02) for b and a estimated parameters, respectively. Average biases were of low magnitudes, ranging from 0.06 (0.06) to 0.07 (0.07) for a and b, and k estimated parameters, respectively (**Table II**).

DISCUSSION

The growth of pigs and differences between individuals' growth curves are controlled by genetic and other effects acting over time (Silva et al., 2017). The genetic analysis of growth curves allows ranking the animals based on their EBVs for the estimates of the growth curve parameters.

Table I. Genetic parameters for Gompertz growth curve parameters in each replicate, as well as averages and respective standard deviations of 10 replicates (Parâmetros genéticos para os parâmetros da curva de crescimento de Gompertz em cada réplica, bem como médias e respectivos desvios-padrão das 10 réplicas).

	Heritability			Genetic correlation		
Replicate	h_a^2	h_b^2	h_k^2	$r_g\hat{a}_i, \hat{b}_i$	$r_g\hat{a}_i, \hat{k}_i$	$r_g\hat{b}_i, \hat{k}_i$
1	0.59 ± 0.07	0.51 ± 0.07	0.51 ± 0.07	0.98 ± 0.01	0.93 ± 0.02	0.95 ± 0.01
2	0.63 ± 0.06	0.58 ± 0.06	0.61 ± 0.06	0.96 ± 0.01	0.92 ± 0.02	0.95 ± 0.01
3	0.62 ± 0.06	0.58 ± 0.06	0.61 ± 0.06	0.97 ± 0.01	0.94 ± 0.01	0.96 ± 0.01
4	0.68 ± 0.06	0.60 ± 0.06	0.55 ± 0.06	0.97 ± 0.01	0.94 ± 0.01	0.95 ± 0.01
5	0.64 ± 0.06	0.59 ± 0.06	0.59 ± 0.06	0.98 ± 0.01	0.95 ± 0.01	0.96 ± 0.01
6	0.58 ± 0.06	0.54 ± 0.06	0.57 ± 0.06	0.97 ± 0.01	0.93 ± 0.02	0.95 ± 0.01
7	0.70 ± 0.05	0.62 ± 0.06	0.61 ± 0.06	0.97 ± 0.01	0.96 ± 0.01	0.97 ± 0.01
8	0.64 ± 0.06	0.55 ± 0.06	0.61 ± 0.06	0.96 ± 0.01	0.93 ± 0.01	0.96 ± 0.01
9	0.58 ± 0.06	0.54 ± 0.06	0.53 ± 0.07	0.95 ± 0.01	0.91 ± 0.02	0.96 ± 0.01
10	0.58 ± 0.06	0.50 ± 0.06	0.55 ± 0.06	0.97 ± 0.01	0.94 ± 0.01	0.96 ± 0.01
Average (SD)	0.62 (0.04)	0.56 (0.04)	0.57 (0.04)	0.97 (0.01)	0.93 (0.01)	0.96 (0.01)

h_a^2: heritability for a estimated parameter (mature body weight); h_b^2 : heritability for b estimated parameter (inflection point); h_k^2: heritability for k estimated parameter (maturation rate); $r_g\hat{a}_i, \hat{b}_i$: genetic correlation between a and b estimated parameters;$r_g\hat{a}_i, \hat{k}_i$: genetic correlation between a and k estimated parameters;$r_g\hat{b}_i, \hat{k}_i$: genetic correlation between b and k estimated parameters; SD: standard deviation.

Table II. Accuracy and bias provided by each replicate, as well as averages and respective standard deviations based on these samples (Acurácia e viés fornecidos por cada réplica, bem como médias e respectivos desvios-padrão baseados nessas amostras).

	Accuracy			Bias		
	Parameters			Parameters		
Replicates	a	b	k	a	b	k
1	0.82	0.81	0.80	-0.09	-0.11	-0.13
2	0.85	0.84	0.84	-0.06	-0.06	-0.03
3	0.85	0.84	0.85	-0.08	-0.06	-0.05
4	0.83	0.81	0.80	-0.01	-0.01	-0.08
5	0.83	0.81	0.82	0.00	0.01	-0.02
6	0.83	0.82	0.83	-0.08	-0.06	-0.05
7	0.87	0.84	0.84	-0.02	-0.04	-0.05
8	0.82	0.79	0.81	0.00	-0.03	0.00
9	0.84	0.83	0.83	-0.19	-0.19	-0.25
10	0.85	0.82	0.83	-0.09	-0.10	-0.07
Average (SD)	0.84 (0.02)	0.82 (0.02)	0.83 (0.02)	0.06 (0.06)	0.06 (0.06)	0.07 (0.07)

a: estimated mature body weight; b: estimated inflection point; k:estimated maturation rate; SD: standard deviation.

In addition, it allows the construction of genetic growth curves for each animal when replacing these EBVs in the Gompertz model equation, which is advantageous, since besides obtaining the EBVs for each time observed, it is possible to estimate these EBVs for any other time of interest within the observed time range. In this sense, the breeding company can detect genetic differences in growth between individuals and outline the best selection strategy to change the shape of animals' growth curves. For that, it is crucial to know the heritability estimates for the growth curve parameters and the genetic correlations between them.

The genetic parameters estimated in the current study (**Table I**) were similar to the parameters used in the simulation, which were based on the results of Cai et al. (2012). The simulated heritabilities were 0.64, 0.58 and 0.60 for a, b and k parameters, respectively, and the simulated genetic correlations were 0.97, 0.94 and 0.96 between a and b, a and k and b and k parameters, respectively. The results show that the two-step method was able to retrieve the simulated genetic parameters. In addition, moderate to high heritability estimates for the growth curve parameters indicate that selection for these parameters could result in reasonable genetic progress through the implementation of these traits in breeding programs (Koivula et al., 2008). The aim of the breeding companies is to change both mature weight and degree of maturity in the same direction, modifying the shape of the growth curves, which may improve the breeding, feeding and management decisions of pig production systems (Lázaro et al., 2017; Niemi, 2006). The high and positive average genetic correlation between a and k estimated parameters found in this study [0.93 (0.01), **Table I**] is favorable for this process, since selection will result in heavier and more precocious animals.

Lázaro et al. (2017) also reported moderate to high heritabilities for a, b and k parameters (0.53, 0.79 and 0.55, respectively) for an outbred F2 (Brazilian Piau x Commercial) pig population and high and positive genetic correlation between a and k parameters (0.78), which is favorable for selection. The heritability estimates of the present study were higher than those reported by Koivula et al. (2008) and Coyne et al. (2017) in pigs, especially for a and k parameters. Differently from the present study, these authors fitted a sire model using a one-step mixed model approach and found heritabilities of 0.44, 0.55 and 0.31 (Koivula et al., 2008) and 0.40, 0.69 and 0.45 (Coyne et al., 2017) for a, b and k parameters, respectively. Coyne et al. (2017) also applied a two-step mixed model approach, and lower heritabilities were found (0.17, 0.32 and 0.18 for a, b and k parameters, respectively). In addition, Koivula et al. (2008) also estimated different genetic correlations compared to the current study, especially between a and k and b and k parameters, with opposite signals (-0.80 and -0.80, respectively), which is not favorable for selection.

The use of simulated data is an effective way to evaluate the prediction process, since the TBVs are known and can be used to validate the EBVs (Oliveira et al., 2019). Our results showed high accuracy and low bias of EBVs, indicating that the two-step process was efficient for the genetic evaluation of repeated growth measurements. In pig breeding programs, longitudinal measurements of individual body weights for group-housed pigs from electronic feeders are becoming increasingly available and important for development of efficient pig breeding strategies (Cai et al., 2012). The present study, based on simulated data, shows an efficient way to deal with these longitudinal data though the fit of a non-linear model and the application of the two-step approach. In addition, it endorses the importance of the knowledge of genetic parameters for the growth curve parameters, since these estimates will directly influence the efficiency of the selection program, especially when they present moderate to high magnitudes, as found in the present study.

CONCLUSIONS

The two-step procedure was able to retrieve the simulated genetic parameters, resulting in heritabilities and genetic correlations similar to the simulated ones. Moderate to high heritability estimates show the possibility of effective selection of pigs based on growth curve parameters. In addition, the high and favorable genetic correlations, especially between the most important growth curve parameters (a and k), show that selection for increased mature weight will also increase maturation rate, which is favorable for pig production systems. High prediction accuracy and low bias estimated values indicate that the two-step approach is feasible for predicting animals' breeding values for growth curve parameters and for dealing with longitudinal body weight measurements.

ACKOWLEGMENTS

The authors would like to thank *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), *Fundação de Amparo à Pesquisa do Estado de Minas Gerais* (FAPEMIG) and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES, grant numbers 88882.306862/2018-01 and PROEX 320.020.170.11P9), for the financial support.

The R codes used in the simulation process are available upon request to Daniele B. D. Marques (danielebdiniz@gmail.com).

BIBLIOGRAPHY

Cai, W, Kaiser, M S, & Dekkers, J C M 2012, 'Bayesian analysis of the effect of selection for residual feed intake on growth and feed intake curves in Yorkshire swine', *Journal of Animal Science*, vol. 90, pp. 127–41.

Coster, A 2012, 'pedigree: Pedigree functions'. R package version 1.4. http://CRAN.R-project.org/package=pedigree.

Coyne, J M, Matilainen, K, Berry, D P, Sevon-Aimonen, M –L, Mäntysaari, E A, Juga, J, Serenius, T, & McHugh, N 2017. 'Estimation of genetic (co) variances of Gompertz growth function parameters in pigs', *Journal of Animal Breeding and Genetics*, vol. 134, pp. 136–43.

Freitas, A R 2005, 'Curvas de crescimento na produção animal', *Revista Brasileira de Zootecnia*, vol. 34, no. 3, pp.786-95.

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

Gompertz, B 1825, 'On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of Life Contingencies', *Philosophical Transactions of the Royal Society of London*, vol. 115, pp. 513–83.

Koivula, M, Sevón-Aimonen, M-L, Strandén, I, Matilainen, K, Serenius, T, Stalder, K J, & Mäntysaari, E A 2008, 'Genetic (co)variances and breeding value estimation of Gompertz growth curve parameters in Finnish Yorkshire boars, gilts and barrows', *Journal of Animal Breeding and Genetics*, vol. 125, pp. 168–75.

Lázaro, S F, Ibáñez-Escriche, N, Varona, L, Silva, F F, Brito, L C, Guimarães, S E F, & Lopes, P S 2017, 'Bayesian analysis of pig growth curves combining pedigree and genomic information', *Livestock Science*, vol. 201, pp. 34–40.

Niemi, J K 2006, 'A dynamic programming model for optimising feeding and slaughter decisions regarding fattening pigs', *Agricultural and Food Science*, Academic Dissertation, vol. 15, pp. 1-121.

Oliveira, H R, Brito, L F, Sargolzaei, M, Silva, F F, Jamrozik, J, Lourenco, D A L, & Schenkel, F S 2019, 'Impact of including information from bulls and their daughters in the training population of multiple-step genomic evaluations in dairy cattle: A simulation study', *Journal of Animal Breeding and Genetics,* vol.136, pp. 1–12.

Pinheiro J, Bates D, DebRoy S, Sarkar D & R Core Team 2015, 'nlme: Linear and Nonlinear Mixed Effects Models', R package version 3.1-120, <URL: http://CRAN.R-project.org/package=nlme>.

R Core Team 2013, 'R: A language and environment for statistical computing', *R Foundation for Statistical Computing*, Vienna, Austria. URL http://www.R-project.org/.

Silva, F F, Resende, M D V, Rocha, G S, Duarte, D A S, Lopes, P S, Brustolini, O J B, Thus, S, Viana, J M S, & Guimarães, S E F 2013, 'Genomic growth curves of an outbred pig population', *Genetics and Molecular Biology*, vol. 36, no. 4, pp. 520-7.

Silva, F F, Zambrano, M F B, Varona, L, Glória, L S, Lopes, P S, Silva, M V G B, Arbex, W, Lázaro, S F, Resende, M D V, Guimarães, S E F 2017, 'Genome association study through nonlinear mixed models revealed new candidate genes for pig growth curves', *Scientia Agricola*, vol. 74, no. 1, pp. 1-7.

Silva, N A M 2008, 'Modelo hierárquico Bayesiano aplicado na avaliação genética de curvas de crescimento de bovinos de corte', Dissertação (Mestrado), Pós-graduação em Estatística e Experimentação Agropecuária, Universidade Federal de Lavras, 47 pp.

Venables, W N & Ripley, B D 2002, 'Modern Applied Statistics with S', Fourth Edition. Springer, New York, ISBN 0-387-95457-0, 516 pp.

Wellock, I J, Emmans, G C, & Kyriazakis, I 2004, 'Describing and predicting potential growth in the pig', *Animal Science*, vol. 78, pp. 379-88.

Wimmer, V, Albrecht, T, Auinger, H J, & Schoen, C C 2012, 'synbreed: a framework for the analysis of genomic prediction data using R', *Bioinformatics*, vol. 28, no. 15, pp. 2086-87.