GENETIC PREDICTION FOR FIRST-SERVICE CONCEPTION RATE IN ANGUS HEIFERS USING A RANDOM REGRESSION MODEL

Miguel A. Sánchez-Castro, Milton G. Thomas, R. Mark Enns, and Scott E. Speidel

Department of Animal Sciences, Colorado State University, Fort Collins, CO 80523-1171

© The Author(s) 2020. Published by Oxford University Press on behalf of the American Society of Animal Science.

INTRODUCTION

First-service conception rate (FSCR) could be defined as the probability of a heifer to conceive in response to her first artificial insemination (AI) and maintain such pregnancy after the end of the breeding season. Economic implications of this trait in beef cattle heifers involve its relationship with their development costs, as well as costs of semen, synchronization protocols, estrus detection and AI services (Horan et al., 2005; Minick Bormann et al., 2006). Furthermore, FSCR is also related to differences in quality and value between AI-produced calves and natural service calves. Additionally, heifers conceiving on their first service, calve earlier within the calving season, have more chances to breed postpartum within a year, and have more time to nurse and wean heavier calves (Marshall et al., 1990).

Given the binary nature of pregnancy outcomes after an AI, the method of choice to evaluate FSCR has been an animal threshold model (ATM). However, such method has a lack of ability to incorporate information of contemporary groups with no variation (e.g., extreme-category problem). When all observations within a categorical fixed effect fall in one of the binary classes, typically, these observations are deleted from the data in order to calculate expected progeny differences (EPD). This can lead to distorted genetic predictions because edited data do not appropriately represent the population’s true variability (Misztal et al., 1989).

Random regression models (RRM) represent an alternative method to evaluate binary traits and can incorporate data from contemporary groups with no variation (Golden et al., 2018). Nonetheless, since they were originally conceptualized to analyze longitudinal traits, the efficacy of RRM to evaluate traits with phenotypes observed only once has not been deeply explored. Relative to heifer fertility, these models have only been applied in genetic evaluations for heifer pregnancy (HPG) in Red Angus cattle (Speidel et al., 2018). Therefore, objectives herein were to evaluate the efficacy of a RRM genetic prediction for heifer FSCR by comparing the resulting EPD to those obtained using a traditional ATM, as well as to compare genetic parameters obtained with each methodology in Angus cattle.

MATERIALS AND METHODS

The data used in the present study were obtained from an existing database; therefore, the study was not subjected to animal care and use committee approval.

Data Collection and Editing

Breeding and ultrasound records of 4,334 Angus heifers (progeny of 354 sires and 1,626 dams) collected from 1992 to 2019 at the Colorado State University Beef Improvement Center
(CSU-BIC) were used for the study. Within each breeding year, heifers were estrus synchronized and subjected to AI only once, before they were exposed to natural service sires approximately 3 wk after insemination. Observations for FSCR (1, successful; 0, unsuccessful) were defined by fetal age obtained from ultrasonic pregnancy exams performed 130 d post-AI. Although FSCR in 12- to 15-mo-old heifers is a single observed phenotype, its expression is likely to be dependent on age of onset of puberty. Consequently, age at AI was an important factor included in all statistical models, and it was calculated as the difference between an individual's birthdate and the date when they were subjected to AI.

Statistical Analysis

Traditional EPD calculation for FSCR was performed using a univariate BLUP threshold animal model along with a probit link function to convert binary observations to an underlying normal distribution. The model equation was as follows:

\[ y^* = X_0 + Z_0 + Q_0 + \epsilon \]

where \( y^* \) corresponded to a vector of transformed observations of FSCR on the underlying scale, \( b \) was a vector of unknown solutions for fixed effects that included breeding contemporary group (defined as a combination between breeding year and semen type), AI technician and age of dam, individual's age at AI was included as a linear covariate, \( u \) corresponded to a vector of unknown solutions of animal random effects, and \( m \) was a vector of unknown solutions of mating group random effects. \( X, Z, \) and \( Q \) were known incidence matrices relating observations in \( y^* \) to both fixed and random effects, and \( e \) was the vector of unknown residual errors. For this model, variances were assumed to be:

\[
\text{Var} \begin{bmatrix} u \\ m \\ e \end{bmatrix} = \begin{bmatrix} A_0 \sigma_a^2 & 0 & 0 \\ 0 & I_n \sigma_m^2 & 0 \\ 0 & 0 & I_n \sigma_e^2 \end{bmatrix}
\]

where \( A_0 \) corresponded to the Wright's numerator relationship matrix, \( I_n \) and \( I_0 \) were identity matrices whose orders were equal to the number of mating groups and observations, respectively. \( \sigma_a^2 \), \( \sigma_m^2 \), and \( \sigma_e^2 \) were the additive, mating group, and residual variances, respectively. In agreement with the specifications of a maximum a posteriori probit threshold model, the residual variance \( \sigma_e^2 \) was constrained to be equal to 1.

Additionally, FSCR was regressed on age at AI using a linear RRM with Legendre polynomials as the base function. The model in matrix form is presented below:

\[ y = Xb + Zu + Qm + \epsilon \]

where \( y \) corresponded to a vector of binary observations of FSCR, \( b \) was a vector of unknown solutions for categorical fixed effects (contemporary group, AI technician, and age of dam) and a linear fixed regression of FSCR on age at AI, \( u \) corresponded to a vector of unknown solutions of animal random additive genetic regression coefficients (intercept and linear), \( m \) was a vector of unknown solutions for mating group random effects. \( X, Z, \) and \( Q \) were known incidence matrices relating observations in \( y \) to both fixed and random effects, and \( e \) was the vector of unknown residual errors. Variances assumed for the models were:

\[
\text{Var} \begin{bmatrix} u \\ m \\ e \end{bmatrix} = \begin{bmatrix} A_0 \otimes G & 0 & 0 \\ 0 & I_n \sigma_m^2 & 0 \\ 0 & 0 & I_n \sigma_e^2 \end{bmatrix}
\]

where \( A \) represented the Wright's numerator relationship matrix, \( \otimes \) was the Kronecker product, and \( G \) corresponds to a modified variance–covariance matrix of additive genetic random regression coefficients where the covariance between the intercept and the linear term was fitted to zero, given no heifer had more than one observation for FSCR. \( I_n, I_0, \sigma_m^2,\) and \( \sigma_e^2 \) remained as described for the previous model. For the estimation of the genetic parameters, the entire pedigree from the CSU-BIC consisting of 13,983 individual animals, as well as 967 and 3,699 unique sires and dams, respectively, was used. The average inbreeding coefficient of this pedigree was 0.009. Once the parameters were obtained, they were compared to each other and subsequently used to estimate EPD for FSCR with each methodology. Resulting predictions were compared through the estimation of Pearson \( (r_p) \) and Spearman’s \( (r) \) correlations. Additionally, a regression coefficient of EPD obtained with the RRM on those obtained with the ATM was estimated. Analyses were performed using ASREML 3.0 (Gilmour et al., 2009) and the Animal Breeder's Tool Kit (Golden et al., 1992).

RESULTS AND DISCUSSION

Summary statistics for FSCR and age at AI are presented in Table 1. Forming contemporary groups by combining breeding year and semen type (i.e., conventional or sexed) resulted in a total of
Genetic evaluations for heifer fertility

Table 1. First-service conception rate and age at first AI summary statistics

| Item                        | N   | Average | SD  | Minimum | Maximum |
|-----------------------------|-----|---------|-----|---------|---------|
| First-service conception rate | 4,334 | 0.46 | 0.50 | 0       | 1       |
| Age at first AI, d          | 4,334 | 422.10 | 21.06 | 347     | 479     |

AI = artificial insemination.

The heritability ($h^2$) estimate of FSCR obtained with the ATM on the underlying scale was 0.03 ± 0.02, which agrees with a previous report in Angus heifers of 0.03 ± 0.03 (Minick Bormann et al., 2006). Even when other reports have indicated FSCR heritability estimates ranging between 0.18 to 0.22, such estimates have been obtained mainly in crossbred cattle (Dearborn et al., 1973; Peters et al., 2013). Conversely, heritability estimates for the resulting intercept and linear term of the RRM were 0.002 ± 0.012 and 0.138 ± 0.078, respectively. Transforming these RRM variance estimates, a $h^2$ of 0.005 ± 0.001 for FSCR at the average age at AI (422 d) was observed. Although the previous estimate was less than that obtained from the ATM, variations in $h^2$ across the range of ages at AI contemplated in this study agree with the previous reports for this trait (Figure 1). Nonetheless, it is important to acknowledge that $h^2$ estimates were considerably higher at the extremes of the age prediction range than in the middle. As a possible explanation for such results, it has been previously reported that a common artifact of RRM using Legendre Polynomials as their base function was their tendency to inflate the genetic variances at the beginning and the end of the covariate data range (Schaeffer and Jamrozik, 2008). This occurs because RRM are sensitive to changes in data distribution, particularly, to reductions in the number of records associated with the covariate implemented (Brügemann et al., 2013). Figure 2 shows the distribution of FSCR records associated with the ages at AI of the Angus heifers from the CSU-BIC. The significant reductions in the number of observations registered at the extremes of the data range could explain the substantial increases in $h^2$ estimates for these ages. Similar data structures have led to comparable variations in $h^2$ estimates for traits like days open and conception rate of dairy cattle when implementing random regression techniques (Yin et al., 2012; Brügemann et al., 2013).

Regarding genetic predictions performed with each methodology, EPD summary statistics are presented in Table 2. Results for the mean EPD were similar between models; however, a wider range in prediction values was observed with the ATM. The smaller variation observed within the
RRM prediction could be explained by the six times smaller \( h^2 \) estimate obtained with this methodology for the average age at AI. A similar outcome was reported by Speidel et al. (2018) when applying RRM in the genetic prediction of HPG in Red Angus cattle. Pearson and Spearman’s correlations among EPD obtained with the ATM and RRM were 0.63 and 0.60, respectively. These results suggested that even when predictions were moderately similar, a considerable re-ranking of individuals occurred between both methodologies and different individuals would be chosen in genetic selection schemes. Sánchez-Castro et al. (2019) reported similar correlation magnitudes (e.g., \( r_p \approx 0.59 \) and \( r_s \approx 0.65 \)) when comparing predictions for stayability in Angus cattle obtained with ATM and RRM. For its part, Speidel et al. (2018) reported higher correlations (e.g., \( r_p = 0.87 \) and \( r_s = 0.89 \)) in a study that compared genetic predictions for HPG obtained with ATM and RRM; however, authors recognized that such results were higher than expected given the differences in the statistical models employed.

Regression of predictions obtained with the RRM on the ATM revealed an underestimation of the genetic merit for FSCR by the RRM in comparison to ATM (\( \beta_1 = 0.095 \)). This difference is likely to be related to the different scale in which the trait is being analyzed within each statistical method (e.g., underlying vs. linear scale). In conclusion, application of RRM for genetic predictions of traits with singly observed phenotypes such as FSCR was feasible. Furthermore, although RRM predictions were moderately similar to those obtained with ATM, the substantial re-ranking of individuals between methodologies suggested that further research is required before considering possible substitutions of threshold evaluations by RRM for use by the beef cattle industry.

**IMPLICATIONS**

This study compiled evidence about the feasibility to apply random regression techniques in genetic evaluations of binary traits like FSCR. Despite the genetic predictions obtained with both methodologies were not perfectly correlated, it is possible that using RRM as opposed to ATM in large-scale genetic predictions for FSCR produce interesting benefits such as the generation of age-specific genetic predictions or increased accuracy of EPD due to their flexibility to include more data. However, regardless of the statistical methodology employed, genetic progress by direct selection on FSCR is expected to be slow due to the low heritability of this reproductive trait.

Table 2. First-service conception rate EPD\(^1\) summary statistics according to the statistical method implemented

| Methodology | \( N \) | Average | SD | Minimum | Maximum |
|-------------|--------|---------|----|---------|---------|
| ATM\(^2\)   | 13,983 | −0.18   | 1.59 | −9.40   | 7.20    |
| RRM\(^3\)   | 13,983 | 0.02    | 0.24 | −1.03   | 1.29    |

\(^1\)EPD = expected progeny differences.
\(^2\)ATM = animal threshold model.
\(^3\)RRM = random regression model.
ACKNOWLEDGMENTS

The authors would like to acknowledge the Mexican National Council for Science and Technology (CONACYT) for providing the scholarship that allowed the first author to conduct research at Colorado State University and to the staff of the John E. Rouse Colorado State University Beef Improvement Center for assistance in collecting data used in this study. This work is supported by USDA National Institute of Food and Agriculture Hatch project COLO0607A, accession number 1006304 and COLO0681A and accession number 1010007. The authors declare that they have no conflict of interest.

LITERATURE CITED

Brügemann, K., E. Gernand, U. U. von Borstel, and S. König. 2013. Application of random regression models to infer the genetic background and phenotypic trajectory of binary conception rate by alterations of temperature×humidity indices. Livest. Sci. 157:389–396. doi:10.1016/j.livsci.2013.08.009

Dearborn, D. D., R. M. Koch, L. V. Cundiff, K. E. Gregory, and G. E. Dickerson. 1973. An analysis of reproductive traits in beef cattle. J. Anim. Sci. 36:1032–1040. doi:10.2527/jas1973.3661032x

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

Golden, B. L., W. M. Snelling, and C. H. Mallinckrodt. 1992. Animal breeder’s toolkit user’s guide and reference manual. Agric. Exp. Stn. Tech. Bull. LTB92-2. Colorado State University, Fort Collins, CO.

Golden, B. L., S. Weerasinghe, B. Crook, S. Sanders, and D. J. Garrick. 2018. A Single-step hybrid marker effects model using random regression for stayability in Hereford cattle. In: Proc. 11th World Congr. Genet. Appl. Livest. Prod., Auckland, New Zealand.

Horan, B., J. F. Mee, P. O’Connor, M. Rath, and P. Dillon. 2005. The effect of strain of Holstein-Friesian cow and feeding system on postpartum ovarian function, animal production and conception rate to first service. Theriogenology 63:950–971. doi:10.1016/j.theriogenology.2004.05.014

Marshall, D. M., W. Mingjiang, and B. A. Freking. 1990. Relative calving date of first-calf heifers as related to production efficiency and subsequent reproductive performance. J. Anim. Sci. 68:1812–1817. doi:10.2527/1990.6871812x

Minick Bormann, J. M., L. R. Totir, S. D. Kachman, R. L. Fernando, and D. E. Wilson. 2006. Pregnancy rate and first-service conception rate in Angus heifers. J. Anim. Sci. 84:2022–2025. doi:10.2527/jas.2005-615

Misztal, I., D. Gianola, and J. L. Foulley. 1989. Computing aspects of a nonlinear method of sire evaluation for categorical data. J. Dairy Sci. 72:1557–1568. doi:10.3168/jds.S0022-0302(89)79267-5

Peters, S. O., K. Kizilkaya, D. J. Garrick, R. L. Fernando, J. M. Reese, R. L. Weaber, G. A. Silver, and M. G. Thomas. 2013. Heritability and Bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. J. Anim. Sci. 91:605–612. doi:10.2527/jas.2012-5580

Sánchez-Castro, M. A., M. G. Thomas, R. M. Enns, and S. E. Speidel. 2019. Stability of genetic predictions for stayability using random regression models that include end points beyond 6 yr of age. Trans. Anim. Sci. 3:1678–1682. doi:10.1093/tas/txz056

Speidel, S. E., R. M. Enns, and B. L. Golden. 2018. Use of a random regression model for the evaluation of Heifer pregnancy in Red Angus cattle. In: Proc. 11th World Congr. Genet. Appl. Livest. Prod., Auckland, New Zealand.

Schaeffer, L. R., and J. Jamrozik. 2008. Random regression models: A longitudinal perspective. J. Anim. Breed. Genet. 125:145–146. doi:10.1111/j.1439-0388.2008.00748.x

Yin, T., B. Bapst, U. U. V. Borstel, H. Simianer, and S. König. 2012. Genetic parameters for Gaussian and categorical traits in organic and low input dairy cattle herds based on random regression methodology. Livest. Sci. 147:159–169. doi:10.1016/j.livsci.2012.04.017