



Comparison of random regression and repeatability models to predict breeding values from test-day records of Norwegian goats

S. Andonov,*¹ J. Ødegård,† M. Svendsen,‡ T. Ådnøy,† M. Vegara,§ and G. Klemetsdal†

*Faculty of Agricultural Sciences and Food, University Ss Cyril and Methodius, PO Box 297, 1000 Skopje, Macedonia

†Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences (UMB), PO Box 5003, N-1432 Ås, Norway

‡Geno, Breeding and AI Association, Norwegian University of Life Sciences (UMB), PO Box 5003, N-1432 Ås, Norway

§Department of International Environment and Development Studies, NORAGRIC, UMB, PO Box 5003, N-1432 Ås, Norway

ABSTRACT

One aim of the research was to challenge a previously selected repeatability model with 2 other repeatability models. The main aim, however, was to evaluate random regression models based on the repeatability model with lowest mean-squared error of prediction, using Legendre polynomials up to third order for both animal additive genetic and permanent environmental effects. The random regression and repeatability models were compared for model fit (using likelihood-ratio testing, Akaike information criterion, and the Bayesian information criterion) and the models' mean-squared errors of prediction, and by cross-validation. Cross-validation was carried out by correlating excluded observations in one data set with the animals' breeding values as predicted from the pedigree only in the remaining data, and vice versa (splitting proportion: 0.492). The data was from primiparous goats in 2 closely tied buck circles (17 flocks) in Norway, with 11,438 records for daily milk yield and 5,686 to 5,896 records for content traits (fat, protein, and lactose percentages). A simple pattern was revealed; for daily milk yield with about 5 records per animal in first lactation, a second-order random regression model should be chosen, whereas for content traits that had only about 3 observations per goat, a first-order polynomial was preferred. The likelihood-ratio test, Akaike information criterion, and mean-squared error of prediction favored more complex models, although the results from the latter and the Bayesian information criterion were in the direction of those obtained with cross-validation. As the correlation from cross-validation was largest with random regression, genetic merit was predicted more accurate with random regression models than with the repeatability model.

Key words: milk trait, model fit, cross-validation

INTRODUCTION

In Norway, a cooperative genetic improvement scheme for buck circles, covering 42% of the national goat population, has been well established (Nævdal et al., 2005). In a buck circle and each year, the participating goat herds (average: 9) typically mate with selected young bucks (5). When their daughters came into production, the sires (2) with the best predicted breeding values were mated to the (30%) elite goats in the circles (Ådnøy et al., 2000). Circles were tied together by use of common elite bucks, either sold live and used in natural mating (after having been used 1 yr in the buck circles where they were proven) or as frozen semen from AI bucks, used across buck circles. In recent years, the number and size of buck circles has been reduced due to a sanitation program for caprine arthritis-encephalitis virus (CAEV), which requires no contact of live animals across herds.

The dairy recording system requires at least 5 test-day records per flock and year for daily milk yield (DMY) of which at least 3 are sampled for fat (FC), protein (PC), and lactose content (LC), both in first and second lactation. Since 1994, the animals have been genetically evaluated by a multitrait repeatability (REP) model, fitting the lactation curve with a fixed Ali-Schaeffer regression (ASR; Ali and Schaeffer, 1987), nested within year-season of kidding (YS), regions, and lactation numbers, and with random effects of flock test day (FTD), animal, and permanent environment (Ådnøy et al., 2000).

Andonov et al. (2007) examined alternatives to the above model (single trait, within region, and in first lactation) and suggested the inclusion of a fixed structure with year (modeled as YS) and a description of the lactation curve by fixed effects of DIM grouped in 3-d periods (DIM3). Here, one aim was to challenge the suggested model with 2 alternative REP specifications: one in which the lactation curve was modeled with ASR, across YS, and another with DIM3, that had flock year (FY) as fixed and FTD as random effects; the latter was hypothesized to increase genetic vari-

Received July 5, 2012.

Accepted November 30, 2012.

¹Corresponding author: sandonov@zf.ukim.edu.mk

ance (Andonov et al., 2007). However, the main aim of the current study was to evaluate random regression (**RR**) models based on the REP model with lowest mean-squared error of prediction, using Legendre polynomials (Kirkpatrick and Heckman, 1989) of different orders for modeling both animal additive genetic and permanent environmental effects, also to estimate genetic parameters for DMY, FC, PC, and LC. Models were compared through model fit, mean-squared error of prediction, and predictive ability by cross-validation.

MATERIALS AND METHODS

Data

The data were the same as used in the study of Andonov et al. (2007), for primiparous goats in 2 closely tied buck circles (7 and 10 flocks, respectively, with 24 to 350 animals per flock over the study period) in one region of Norway (Hedmark). Data was from years 1988 to 2003, with records for DMY ($n = 11,438$ from 2,371 goats), FC and PC ($n = 5,896$ from 2,215 goats), and LC ($n = 5,686$ from 2,111 goats). Records from 5 to 305 DIM were kept when ranging 0.1 to 9.9 kg per day for DMY. Fat content, PC, and LC values between 1 and 9% were accepted. Each year, 3 seasons of kidding were used (December to February, March to May, and June to November, with 1 to 464 records for content traits), in accordance with the current practice in genetic evaluation.

The pedigree data involved all animals with records and their ancestors traced back as far as possible. The pedigree data contained 5,511 animals (75.96% had both parents known); of these, 421 were sires and 5,090 were dams, with 261 animals defining the base population. For animals with DMY records, the average number of generations known [the complete generation equivalent, according to Boichard et al. (1997)] was 3.27.

REP Models

One aim of the study was to challenge the preferred model of the previous study (Andonov et al., 2007) defined as model 1:

$$y_{ijkl} = YS_i + DIM\beta_j + FTD_k + a_l + pe_l + e_{ijkl}$$

where y_{ijkl} is DMY, FC, PC, or LC for goat l within YS class i , stage of lactation j , and FTD k ; YS_i is fixed YS effect, with 42, 41, 41, and 36 classes for DMY, FC, PC and LC, respectively; $DIM\beta_j$ is a fixed effect of stage of lactation, defined in 3-d intervals starting from d 5, with 101 classes; FTD_k is the random effect of FTD k , whereas the random animal additive genetic effect of

goat l is a_l ; and pe_l and e_{ijkl} are random effects of permanent environment of animal and residual, respectively.

The competing model had ASR across YS (model 2):

$$y_{ikl} = YS_i + b_1(DIM/305) + b_2(DIM/305)^2 + b_3 \ln(305/DIM) + b_4 [\ln(305/DIM)]^2 + FTD_k + a_l + pe_l + e_{ikl}$$

where additional notation relative to that described for model 1 is b_1 , b_2 , b_3 , and b_4 , which are fixed ASR coefficients across YS classes, and DIM, which is day of lactation (5 to 305).

Finally, another specification capable of accounting for time trends was considered by omitting YS in model 1 and rather considering the effect of the m th FY as a fixed effect (FY_m), otherwise for comparable notation (model 3):

$$y_{jklm} = DIM\beta_j + FY_m + FTD_k + a_l + pe_l + e_{jklm}$$

As DMY was recorded more frequently than milk content in a flock, the number of observations per level of FY was higher for DMY than for FC, PC and LC (155, 146, 146 and 134 classes, with average number of records 73.8, 40.4, 40.4 and 42.4, respectively).

RR Models

The best (see Results and Discussion section) REP model (model 1) was extended by adding Legendre polynomials of either first, second, or third order, both for the random effects of animal and the permanent environment, resulting in models denoted as **RRL1**, **RRL2**, and **RRL3**, respectively. The RRL models can be described as follows:

$$y_{ijklo} = YS_i + DIM\beta_j + FTD_k + \sum_{m=0}^3 a_{lm} Z_{om} + \sum_{m=0}^3 pe_{lm} Z_{om} + e_{ijklo}$$

where the additional notation to that in model 1 is Z_{om} , the polynomial m for DIM o , where $m = 0, 1, 2$, or 3, of same order for both the animal and permanent environmental effects; a_{lm} is the random regression coefficient on Z_{om} for the animal genetic effect; pe_{lm} is the random regression coefficient on Z_{om} for the permanent environmental effect; and e_{ijklo} is the random residual.

At DIM o , additive genetic variance ($\sigma_{a_o}^2$), permanent environmental variance ($\sigma_{pe_o}^2$), and heritability (h_o^2) were calculated as follows:

Download English Version:

<https://daneshyari.com/en/article/10976233>

Download Persian Version:

<https://daneshyari.com/article/10976233>

[Daneshyari.com](https://daneshyari.com)