Estimates of Variance Components for Test-Day Models with Legendre Polynomials and Linear Splines

J. Bohmanova[1], J. Jamrozik[1], F. Miglior[12], I. Misztal[4]

[1]CGIL, University of Guelph, Guelph, ON, [2]Agriculture and Agri-Food Canada, Sherbrooke, QC,
[3]Canadian Dairy Network, Guelph, ON, [4]University of Georgia, Athens, GA, USA

Agriculture and
Agri-Food Canada

Agriculture et
Agroalimentaire Canada

Introduction

In the current Canadian Test-Day Model, Legendre polynomials of order four are fitted for both fixed and random regressions. High additive genetic variances at extremes of lactations has been reported in random regression models (RRM) based on Legendre polynomials. Recently, spline functions has been advocated as a good alternative to Legendre polynomials, due to theirs higher flexibility and limited sensitivity to the data.

Objectives

Compare currently used RRM based on Legendre polynomials with four models based on linear splines using visual inspection of variance function and Deviance Information Criterion.

Materials and methods

Data

- ✧ 96,756 test-day records of 6,094 Canadian Holstein cows
- ✧ 4 traits - milk, fat and protein yields, and somatic cell score (SCS)
- ✧ 3 lactations, pedigree file contained 18,178 animals.

Models

The general formula for the four competing random regression models was:

$$y_{ijkmnt} = HTD_{jkn} + \sum_{l=1}^{q} \alpha_{jlmn} z_l(d) + \sum_{l=1}^{q} \beta_{ijln} z_l(d) + \sum_{l=1}^{q} \gamma_{ijln} z_l(d) + e_{ijkmnt}$$

- ✧ HTD = herd-test-date effect
- ✧ α = fixed regression on DIM x age-season-region of calving class
- ✧ β = random regressions on DIM for additive genetic effect
- ✧ γ = random regressions on DIM for permanent environmental effect

The fixed and random regressions were fitted either with:

- ✧ Legendre polynomials of order four (**LEG**)
- ✧ linear splines with four (**SPL4**), five (**SPL5**) or six (**SPL6**) knots

Model	Type of regression function	Number of knots	Position of knots
LEG	Legendre	5	-
SPL4	Splines	4	[5 125 245 365]
SPL5	Splines	5	[5 65 125 245 365]
SPL6	Splines	6	[5 65 125 245 305 365]

Parameters were estimated with a Bayesian approach via Gibbs sampling

- ✧ a single chain of 100,000 samples
- ✧ 20,000 samples discarded as a burn-in

Model comparison

For each model the Deviance Information Criterion (DIC) defined as:

$DIC = \bar{D} + p_D$ was calculated, where:

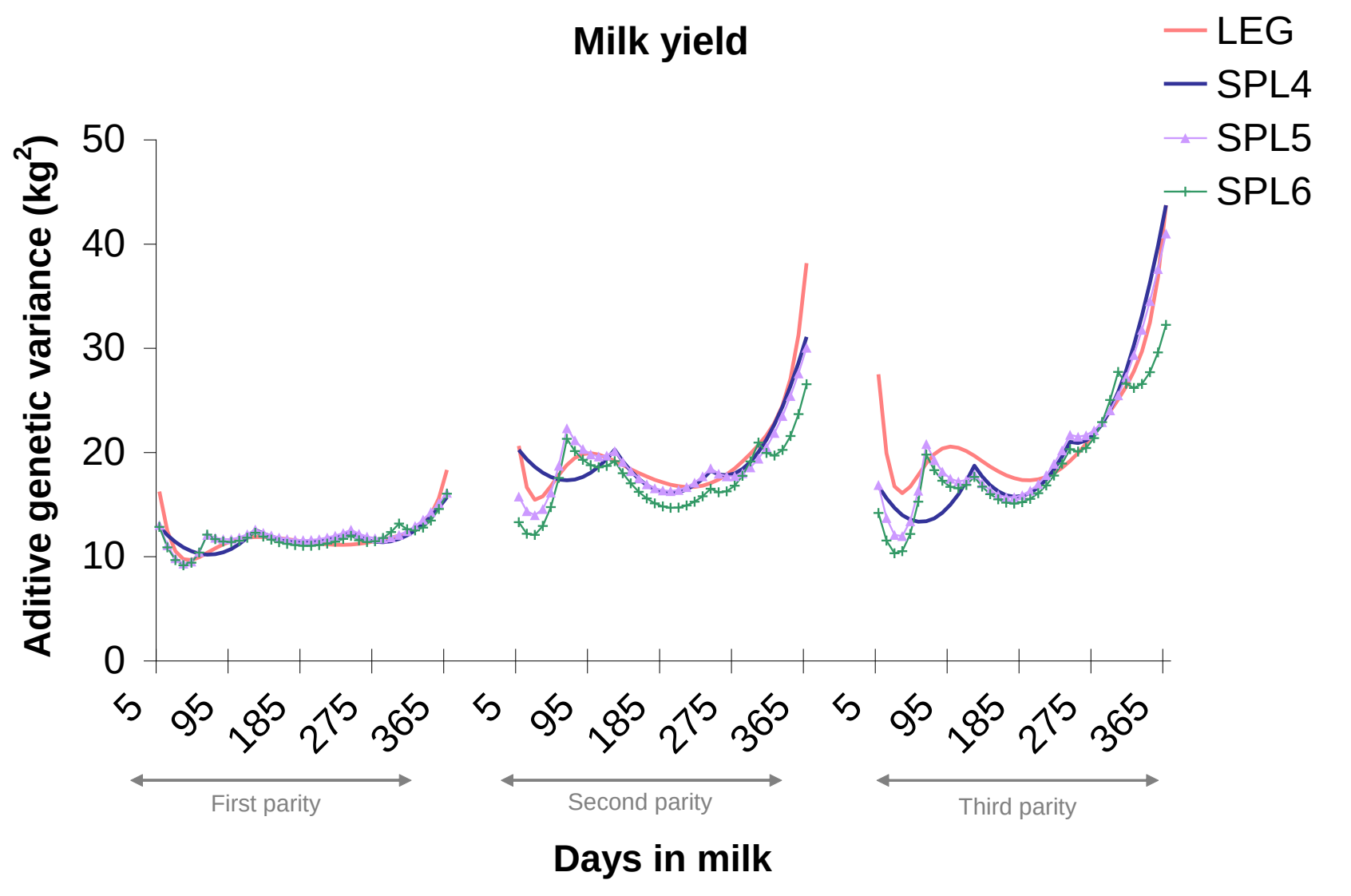
- $\bar{D}$ is the posterior expectation of the Bayesian deviance (measure of the fit of the model)
- p_D is the effective number of parameters (penalty for increasing model complexity),

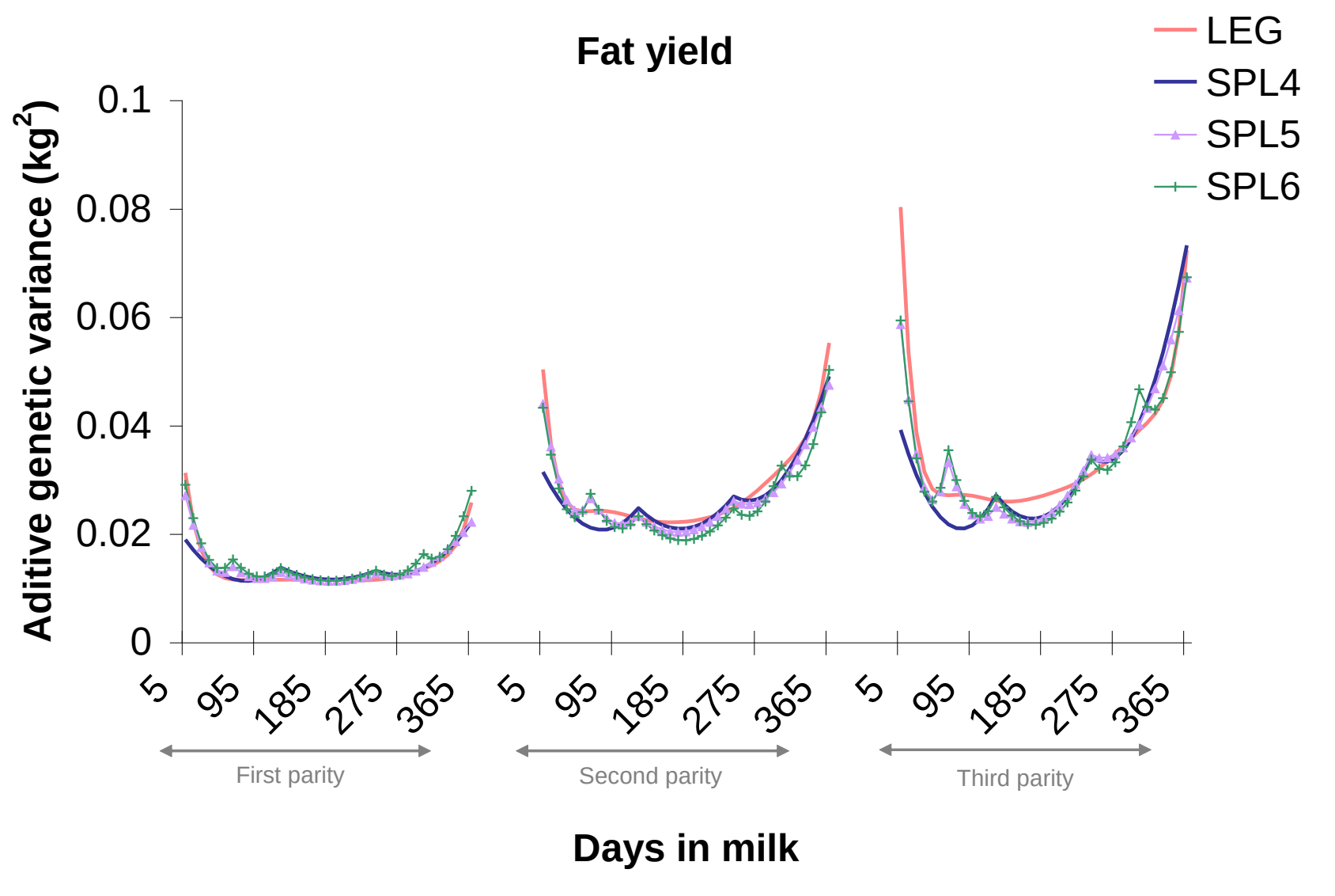
The model with the smallest DIC was preferable.

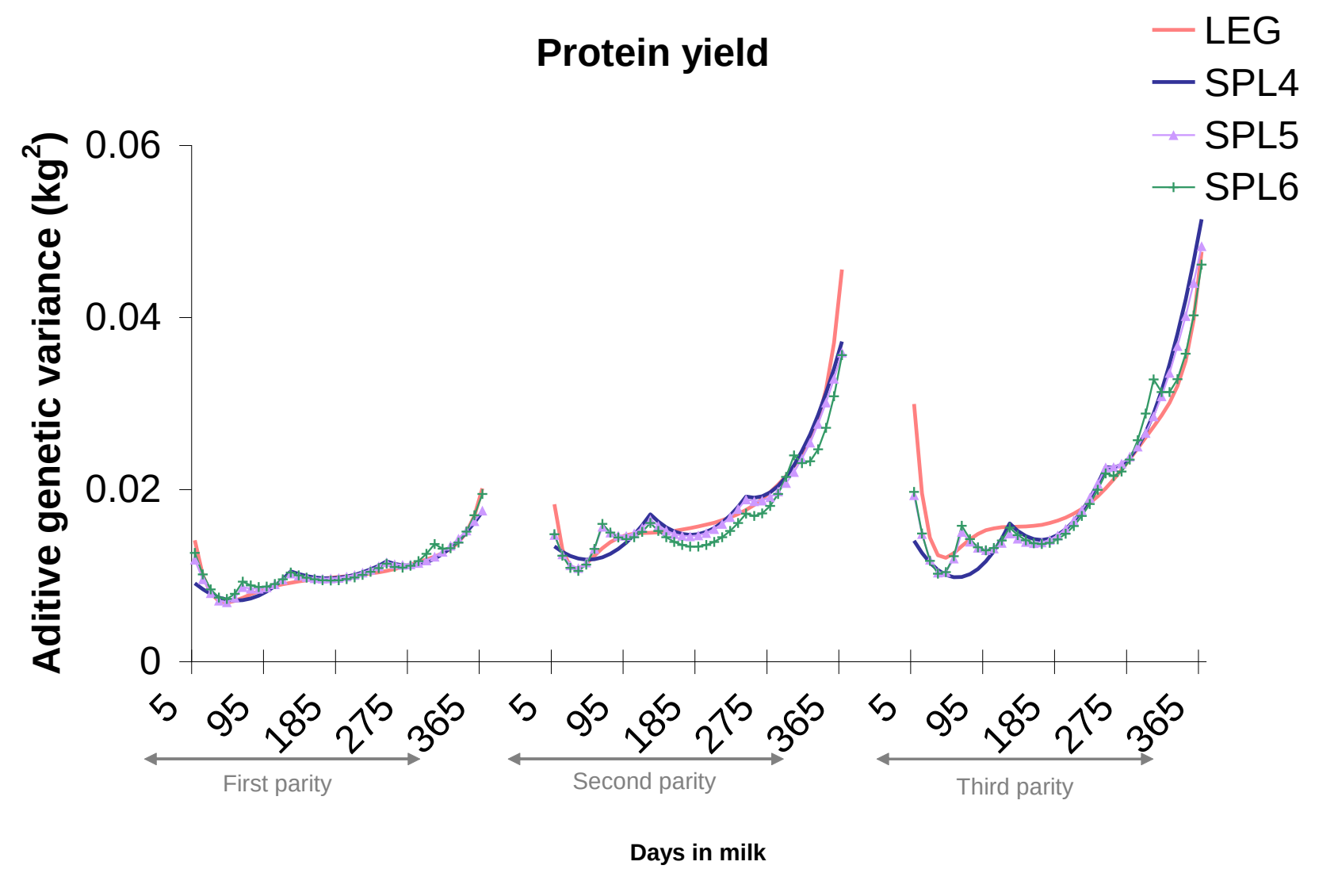
Results

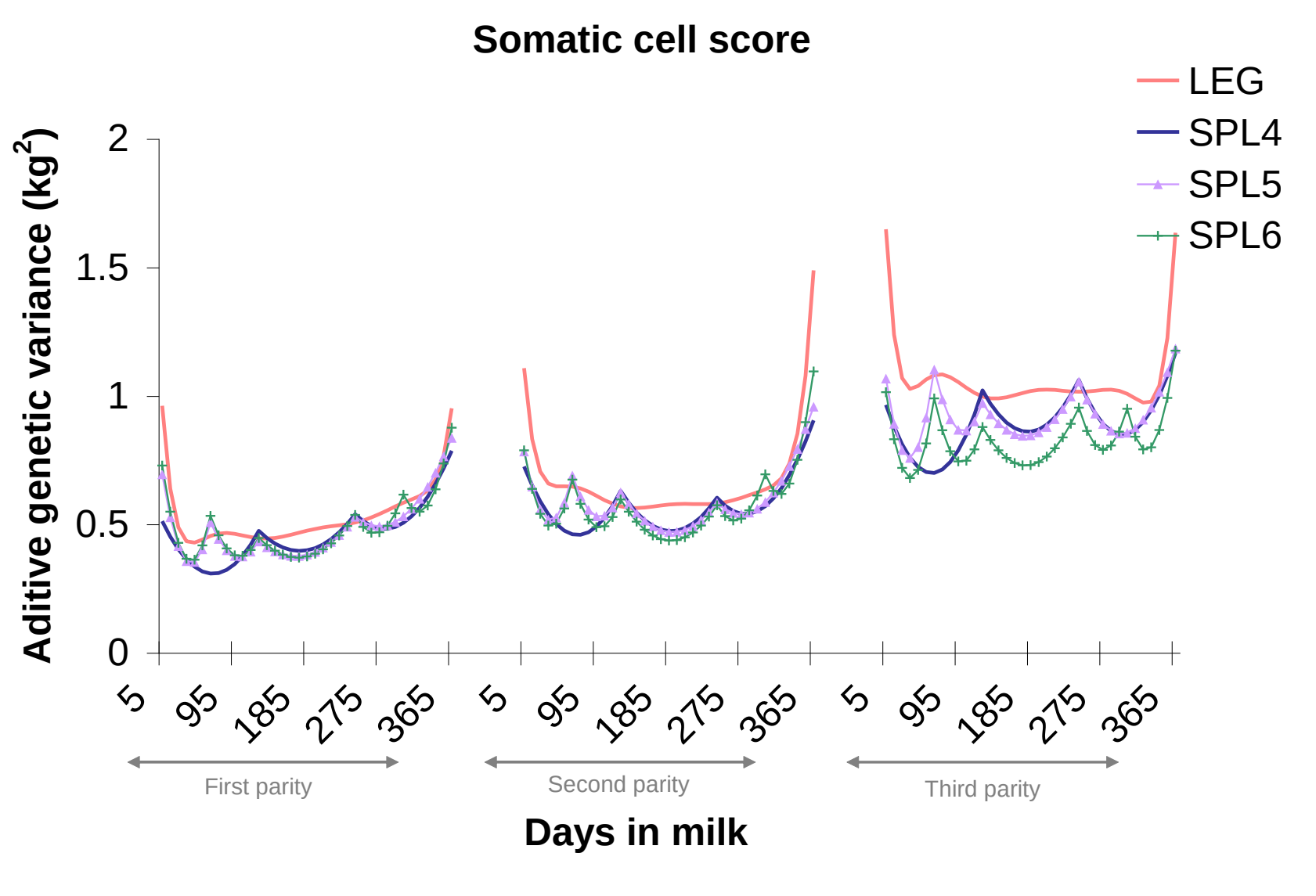
In terms of DIC, the SPL6 was superior to other models in both goodness of fit and model parsimony.

Model	DIC	DIC/DIC_{LEG}	Rank of models
LEG	255,808	1.00	2
SPL4	270,055	1.05	4
SPL5	258,924	1.01	3
SPL6	236,646	0.93	1









Conclusions

The SPL6 model had a better goodness of fit than currently used LEG model. Smaller additive genetic variances were found at extremes of lactations in models with splines compared to the model with Legendre polynomials.

Acknowledgements

Funding was provided by DairyGen Council of Canadian Dairy Network and NSERC of Canada.

Canada