56th Annual Meeting of the European Association for Animal Production June 5-8 2005, Uppsala, Sweden

Modelling repeated measures of ejaculate volume of Holstein bulls using a Bayesian random regression approach

M. Serrano, M.J. Carabaño and C. Díaz

Departamento de Mejora Genética Animal,INIA
28040 Madrid, Spain

- **In dairy cattle breeding, increasing interest in functional traits**

male fertility ←→ semen characteristics

- **AI centers routinely collect data on semen production (quantity and quality).**

Repeated measures along time

- **Alternative models are possible.**

Simple repeatability ←←-------→→ Multiple trait

Random regression

OBJECTIVE

- Use of alternative random regression models to analyse measures of first ejaculate volume of Holstein bulls.

DATA

- 8773 weekly records between 12 and 30 mo. of age of
- 213 Holstein bulls.→ 57.2 measures/ bull (8-92)
- 1140 animals in pedigree

STATISTICAL ANALYSIS - MODELS

RANDOM REGRESSION MODELS (RRM)

$$y_{ijkl} = MY_i + YF_j + \sum_{m=1}^{p} \beta_{jm} P_{X_m}(t) + \sum_{m=0}^{p} \alpha_{km} P_{Z_{km}}(t) \sum_{m=0}^{p} \omega_{lm} P_{W_{lm}}(t) + + e$$

Linear RRM

$$Y_{ijkl} = MY_i + YF_j + \beta_{1j} t + \alpha_{0k} + \alpha_{1k} t + \omega_{01} + \omega_{11} t + e_{ijkl}$$

- MY= month-year of collection; YF= year of first collection
- β, α, ω = regression coefficients for fixed environmental, additive genetic and permanent environmental components
- _P_X, P_Z, P_W = Corresponding Legendre polynomial covariates
- e_{ijkl} = residual component

STATISTICAL ANALYSIS - MODELS

- Alternative models,
  - Polynomials order (p=1,...,6→ LEG-1,...,LEG-6)
  - t= Age at collection (AC) vs. days in production (DP)
  - Homogeneous (HOM) vs. Heterogeneous (HET) residual variance

6 x 2 x 2 = 24 models analysed

STATISTICAL ANALYSIS – Bayesian Inference

Inference on the unknowns in <u>EACH MODEL</u> provide,

- EBV for regression coefficients

$\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_p$ Biological interpretation ???

- Genetic and permanent environmental (co)variances for the regression coefficients (pxp matrices) ???

From those values, can infer the expected trajectories of

- EBV at each time t (AC or DP)
- variances (heritabilities) in each time t

STATISTICAL ANALYSIS – Bayesian criteria for model comparison

- **Log-Marginal Density of Data(<u>LMD</u>)–"Goodness of fit"**

- **Cross validation predictive densities**

$$\underline{D} = \frac{1}{n}\sum_{i=1}^{n} d_r^2 \qquad , \text{ with } d_r = E_{Y_r \mid y(r)} \; [y_r - Y_r]$$

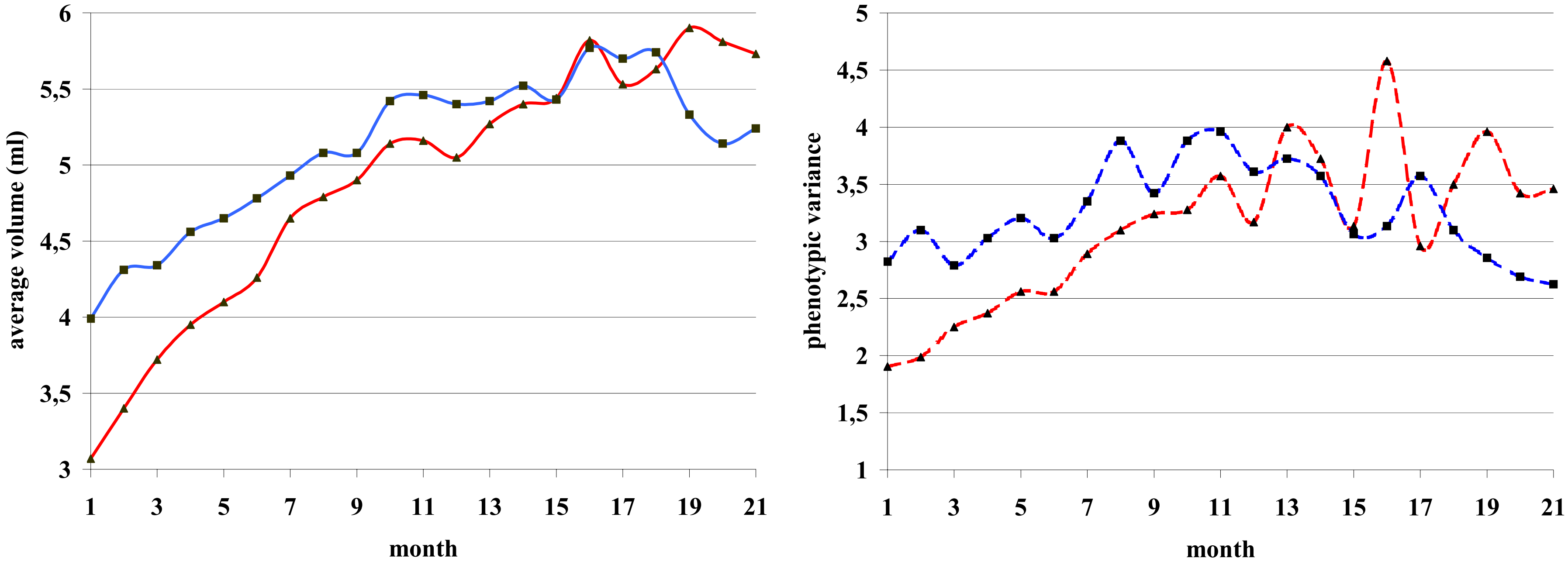
prediction

observation

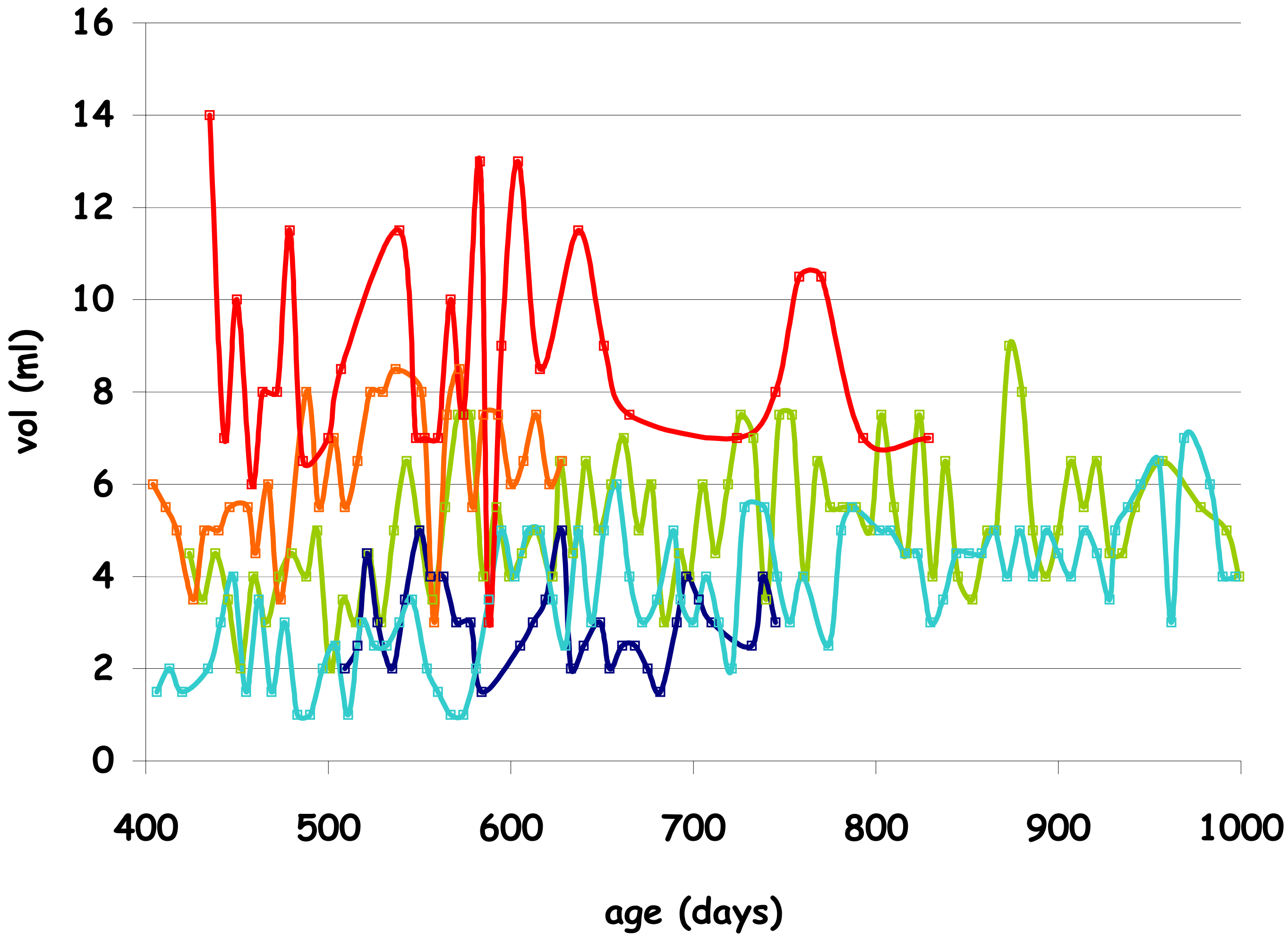
data except Y_r

PHENOTYPIC DATA -- AVERAGES

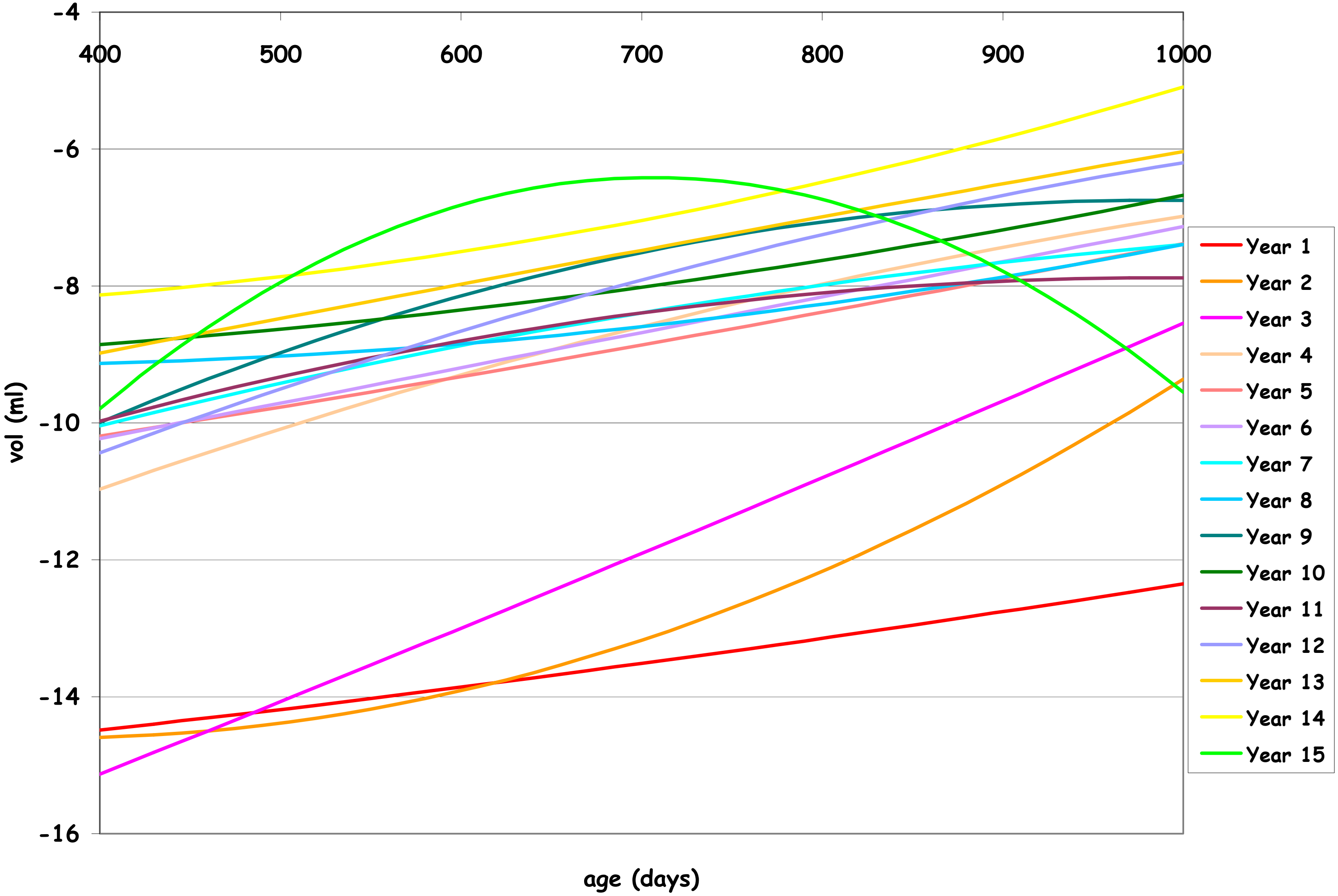
Mean (solid) and variance (dash) of semen volume in first ejaculate at different months of age (-▲-) and months in production (-■-).



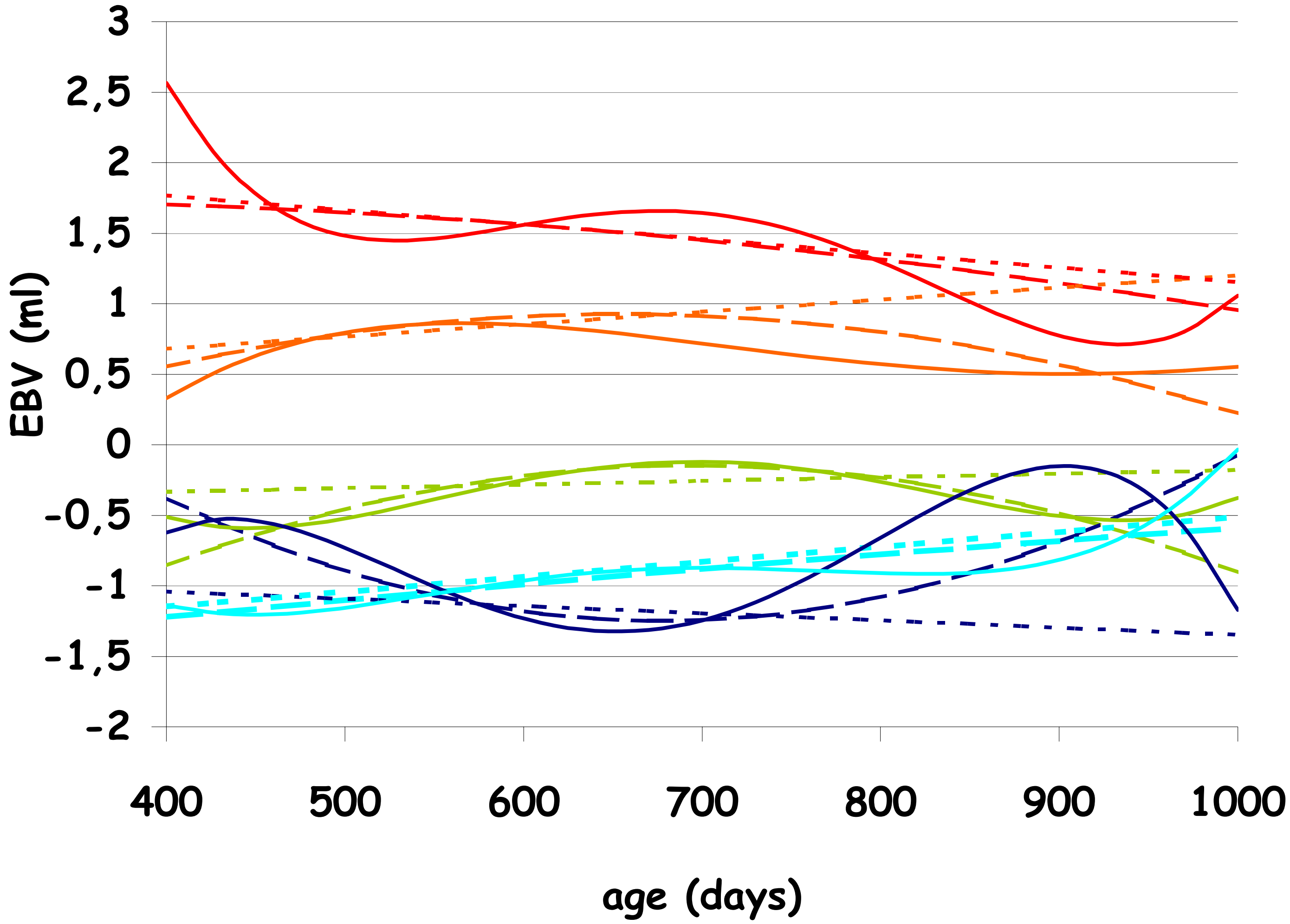
PHENOTYPIC DATA – 5 INDIVIDUALS



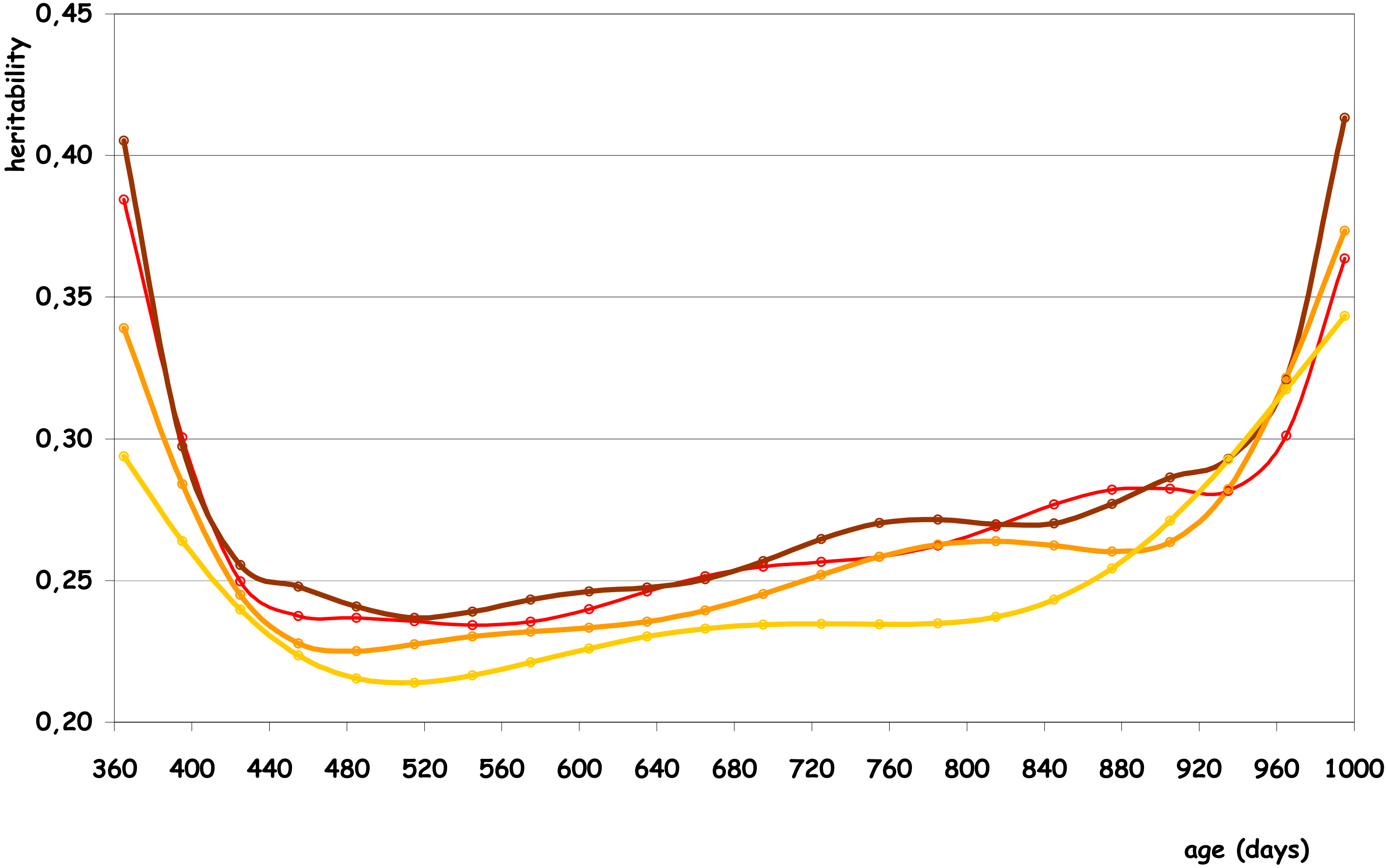
Estimated average semen production curves nested to year of first collection (t=AC; LEG-2)



Estimated EBVs of five bulls at different ages LEG-1 (.....), LEG-2 (_ _ _) and LEG-4 (____)

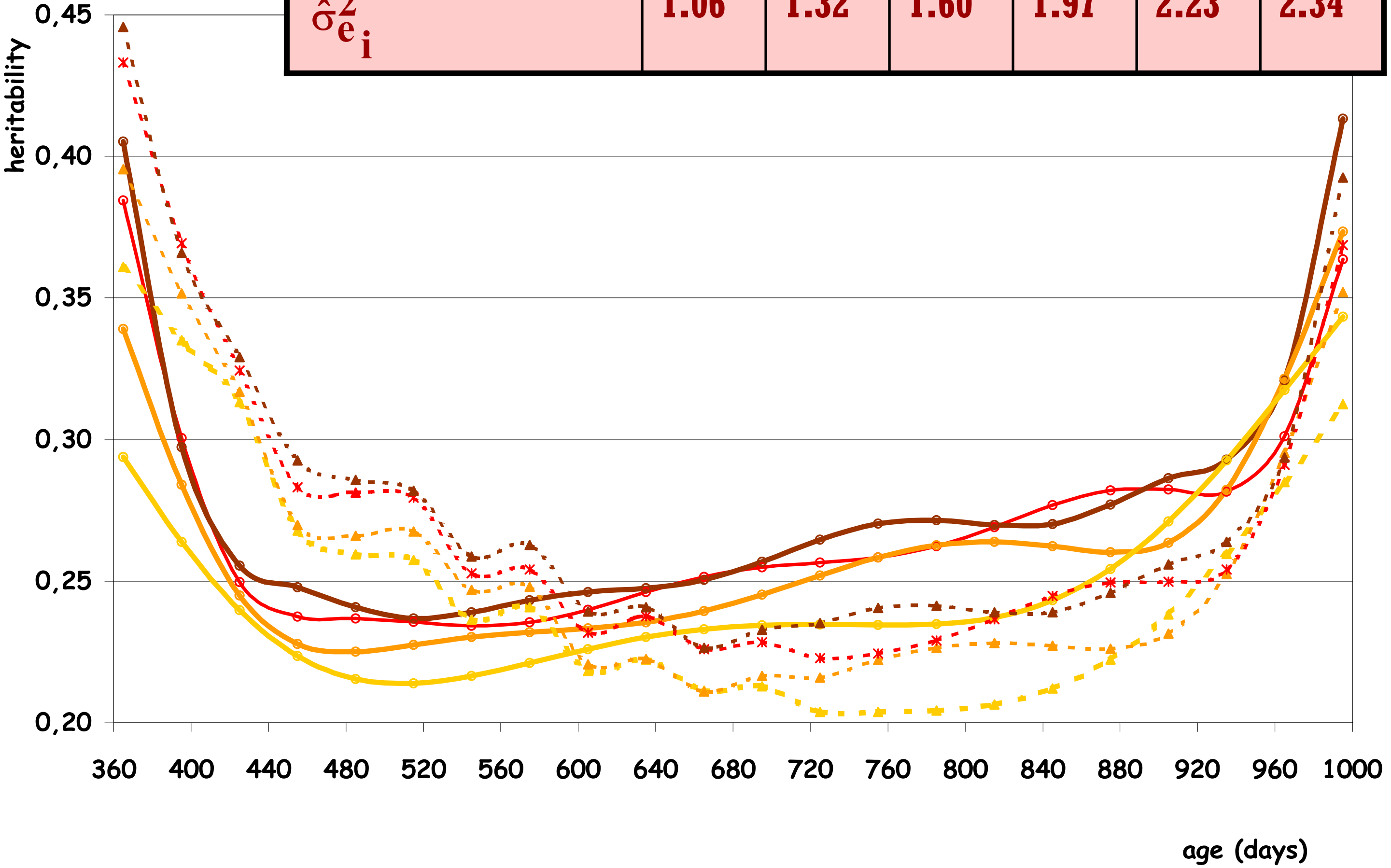


Estimated heritabilities along time (t=AC) and different ordes of polynomials (LEG-2, LEG-3, LEG-4, LEG-5) under HOM residual variance



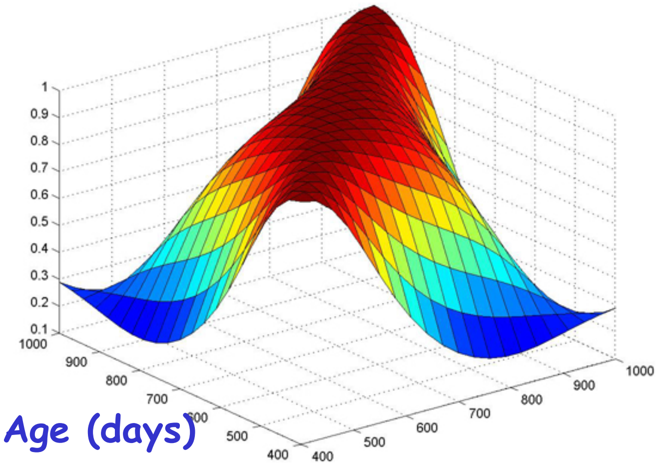
Estimated residual variances ($\hat{\sigma}^2_{e_i}$) and heritabilities ……. under HOM(solid) or HET (dotted) residual variance

Upper limit (days)	430	520	580	640	700	1000
$\hat{\sigma}^2_{e_i}$	1.06	1.32	1.60	1.97	2.23	2.34

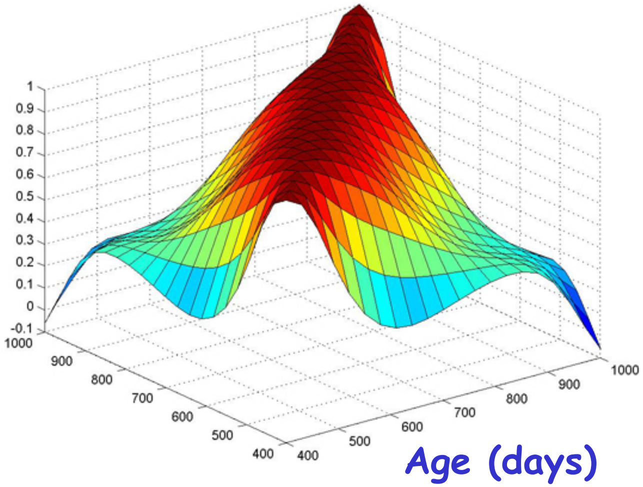


GENETIC CORRELATIONS – t=AC

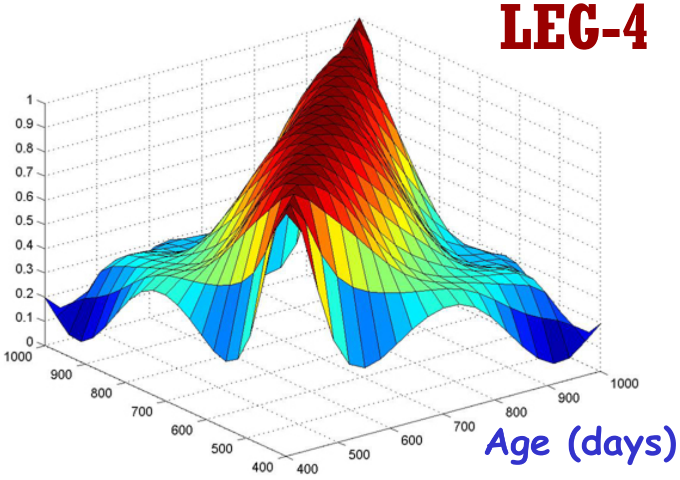
LEG-2



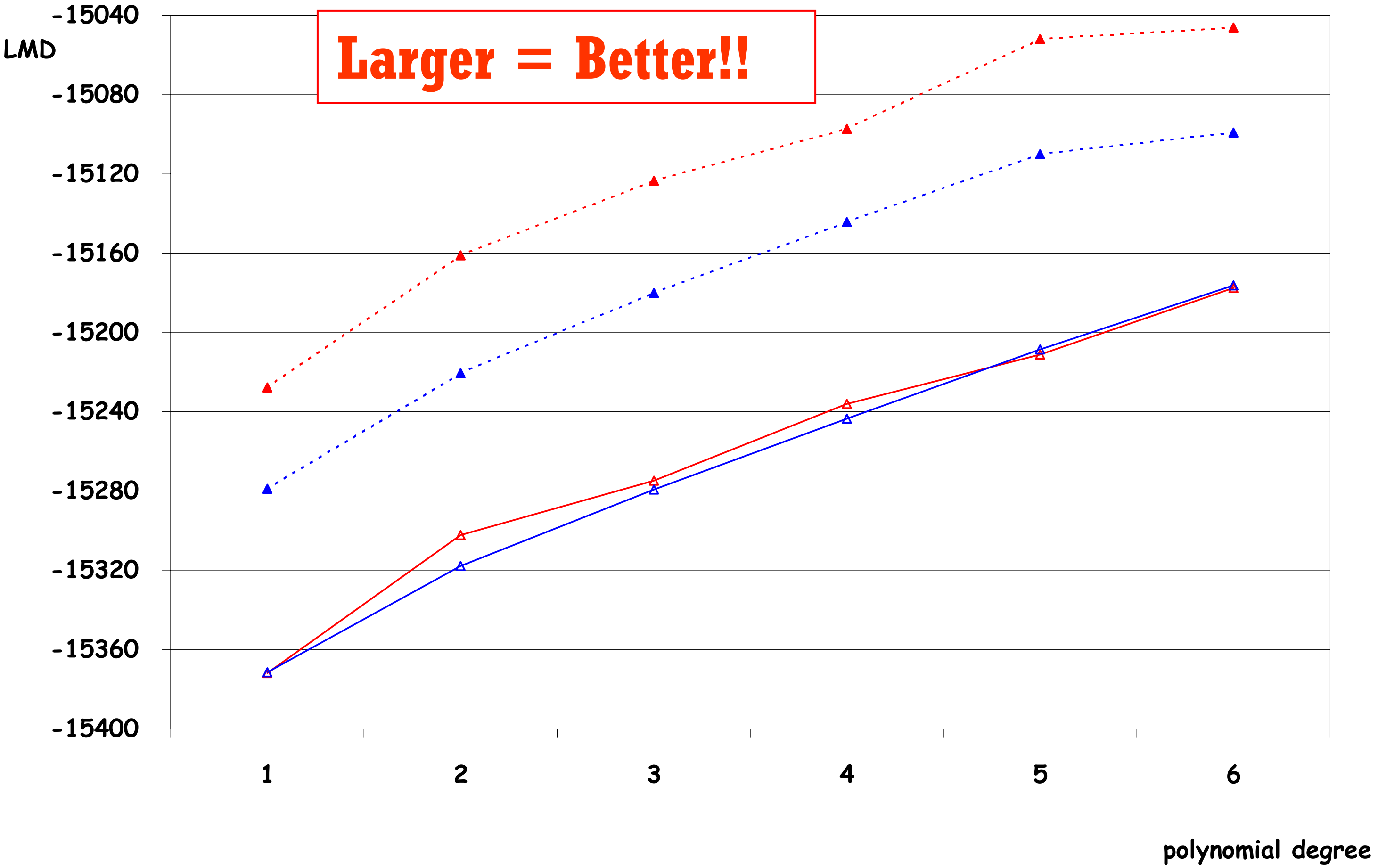
LEG-3



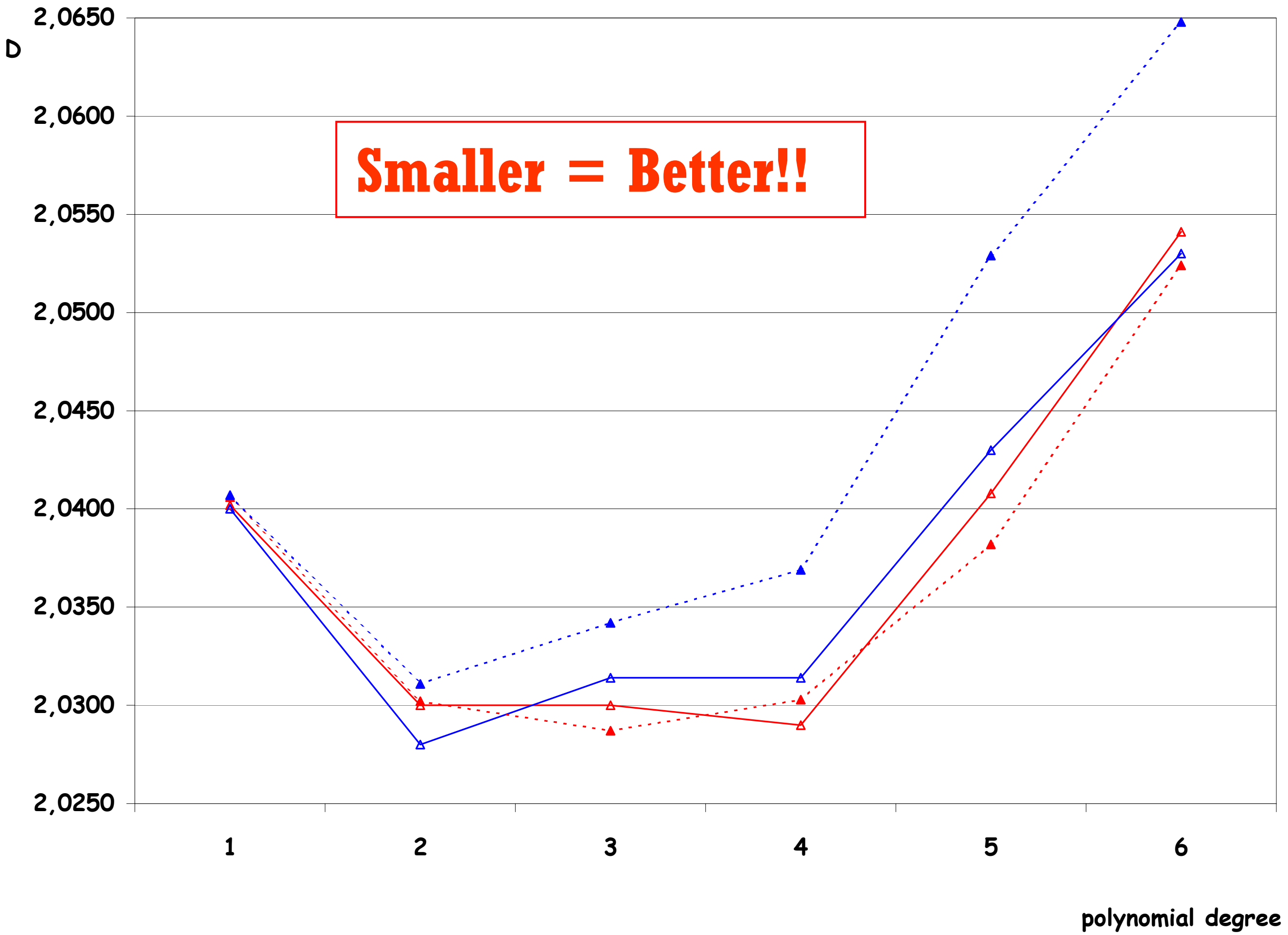
LEG-4



LMD for models differing in order of polynomial for t=AC (-▲-) or t= DP (-■-) under HOM (solid0 or HET (dotted) residual variance



Cross validation D statistics for models differing in order of polynomial for t=AC (-▲-) or t= DP (-■-) under HOM (solid0 or HET (dotted) residual variance



CONCLUSIONS

- The genetic potential to produce semen changes with age existing a variety of patterns among individuals.

- Volume in first ejaculate shows a moderate heritability (0.2 to 0.3, ignoring edges) indicating that selection for this trait may be effective at any stage.

- RRM are a useful tool to analyse semen characteristics because they provide information on both the level of production and its changes along time. However, care has to be taken when interpreting results at the extremes of the period.

- A Legendre polynomial of order 3-4 using t=AC and HET residual variance would be recommended for this data set.

AKNOWLEDGEMENTS

- The authors thank ABEREKIN, S.A. for providing the data on semen traits and CONAFE for providing genealogical information.

STATISTICAL ANALYSIS – Bayesian Inference

- **Sampling distribution**

$y \mid b, \alpha, \omega, R \sim MVN(Xb + Z\alpha + W\omega, R), R=I\sigma_{ei}^2, i = 1, \ldots, \delta$

- **Prior distributions for the unknowns**

Regression coefficients, $\beta, \alpha, \omega \sim$ **MVN**

Residual variance $\sigma_{ei}^2 \sim \chi^{-2}$

Genetic, Permanent environmental (co)variances of RC ~ **IW**

- **Gibbs sampling – Coupled chains (10,000 burn-in/ 120,000 total)**

STATISTICAL ANALYSIS
EBV and variances along time

- **Additive genetic deviation for animal i at time t,**

$$a_i(t) = z\,(t)'\alpha_i$$

- **Additive genetic (G) and permanent environmental (P) (co)variances,**

$$G(t,\ t') = Z(t)\Sigma_a\ Z(t')'$$

$$P(t,\ t') = W(t)\ \Sigma_p\ W(t')$$

GENETIC VARIANCES -- AC

Estimates of genetic variances for <u>AGE AT COLLECTION</u> (AC) und models of increasing polynomial order