Test day models

Vincent Ducrocq

**(+ Hélène Leclerc, <u>Tom Druet</u>,
Pilar Schneider, Clotilde Patry, Mathieu Arnal)**

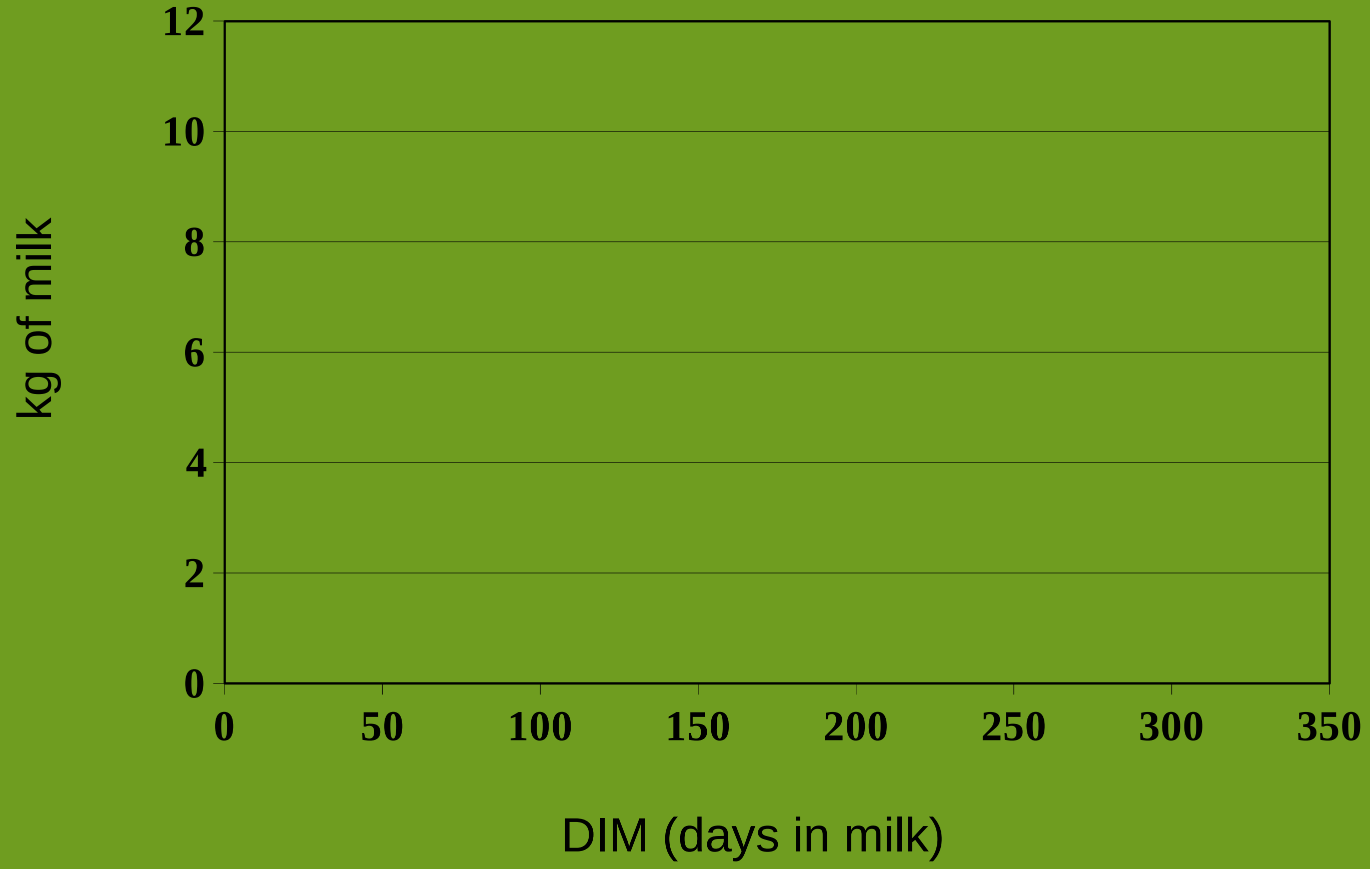
GABI, Jouy-en-Josas

INRA
SCIENCE & IMPACT

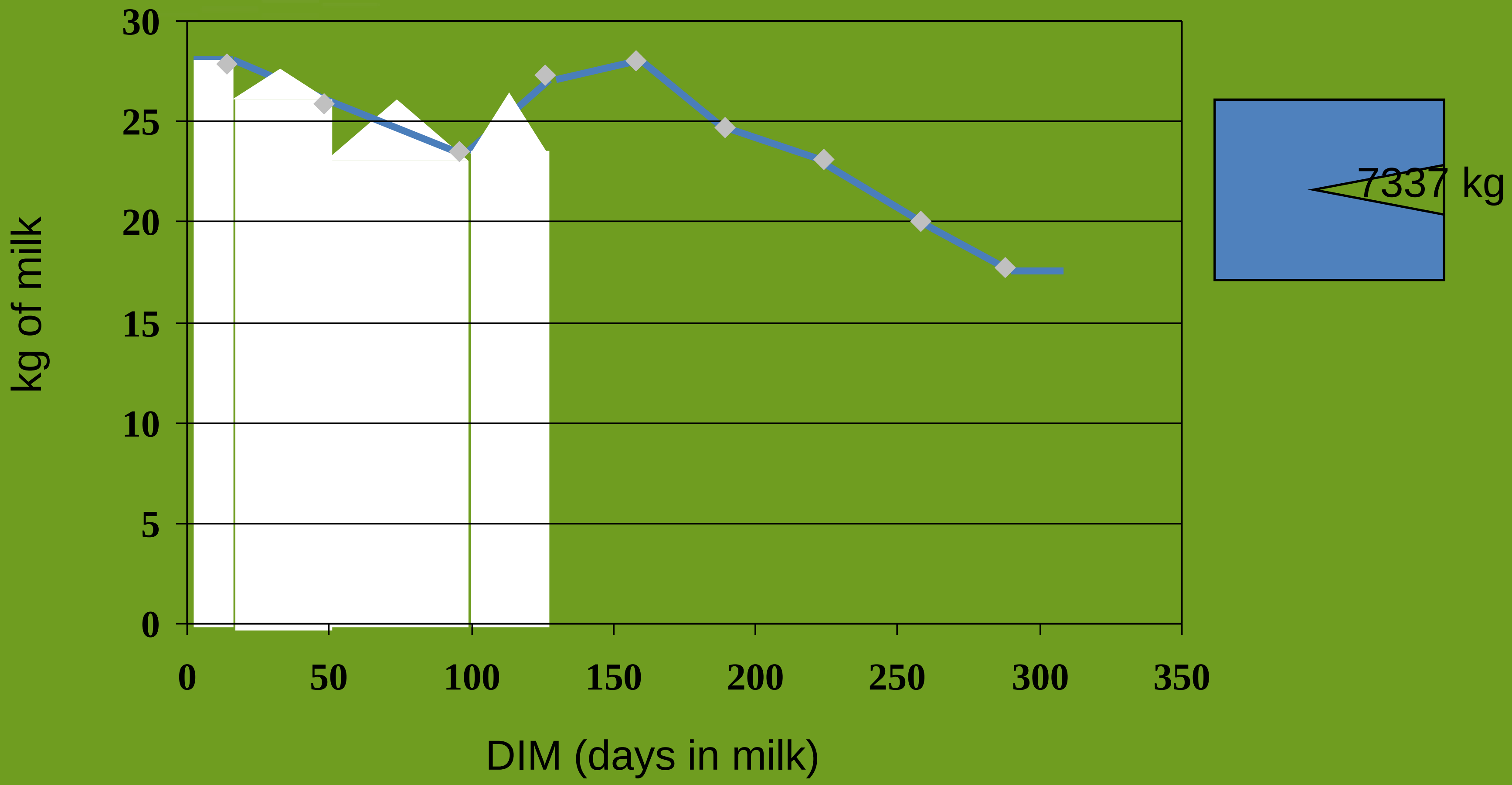
Data from one lactation

Calving date	*Test date*	*Days in milk (DIM)*	*Milk production (kg)*
03/10/2010	*17/10/2010*	*14*	*27.9*
03/10/2010	*20/11/2010*	*48*	*25.9*
03/10/2010	*07/01/2011*	*96*	*23.5*
03/10/2010	*06/02/2011*	*126*	*27.3*
03/10/2010	*10/03/2011*	*158*	*28.0*
03/10/2010	*10/04/2011*	*189*	*24.7*
03/10/2010	*15/05/2011*	*224*	*23.1*
03/10/2010	*18/06/2011*	*258*	*20.0*
03/10/2010	*18/07/2011*	*288*	*17.7*

Example of one lactation

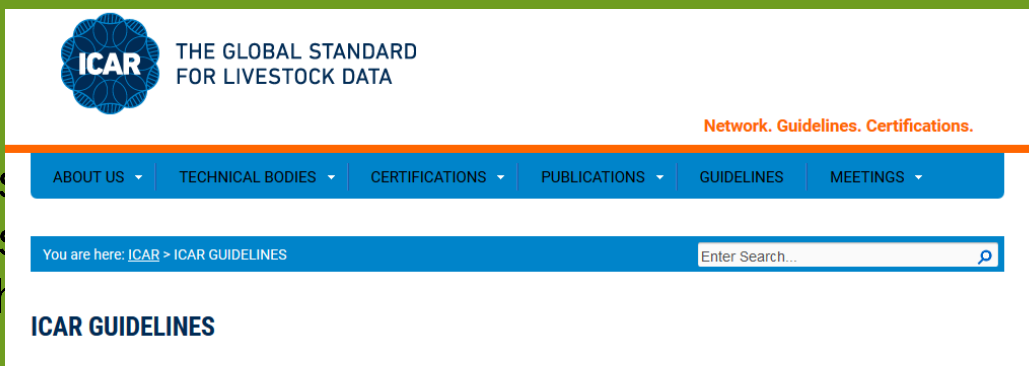


Example of one lactation



How to compute a lactation record?

- ICAR (*International Committee for Animal Recording*) has developed official procedures for recording nearly any trait in animals
- *http://www.icar.org/index.php/icar-recording-guidelines/*



S

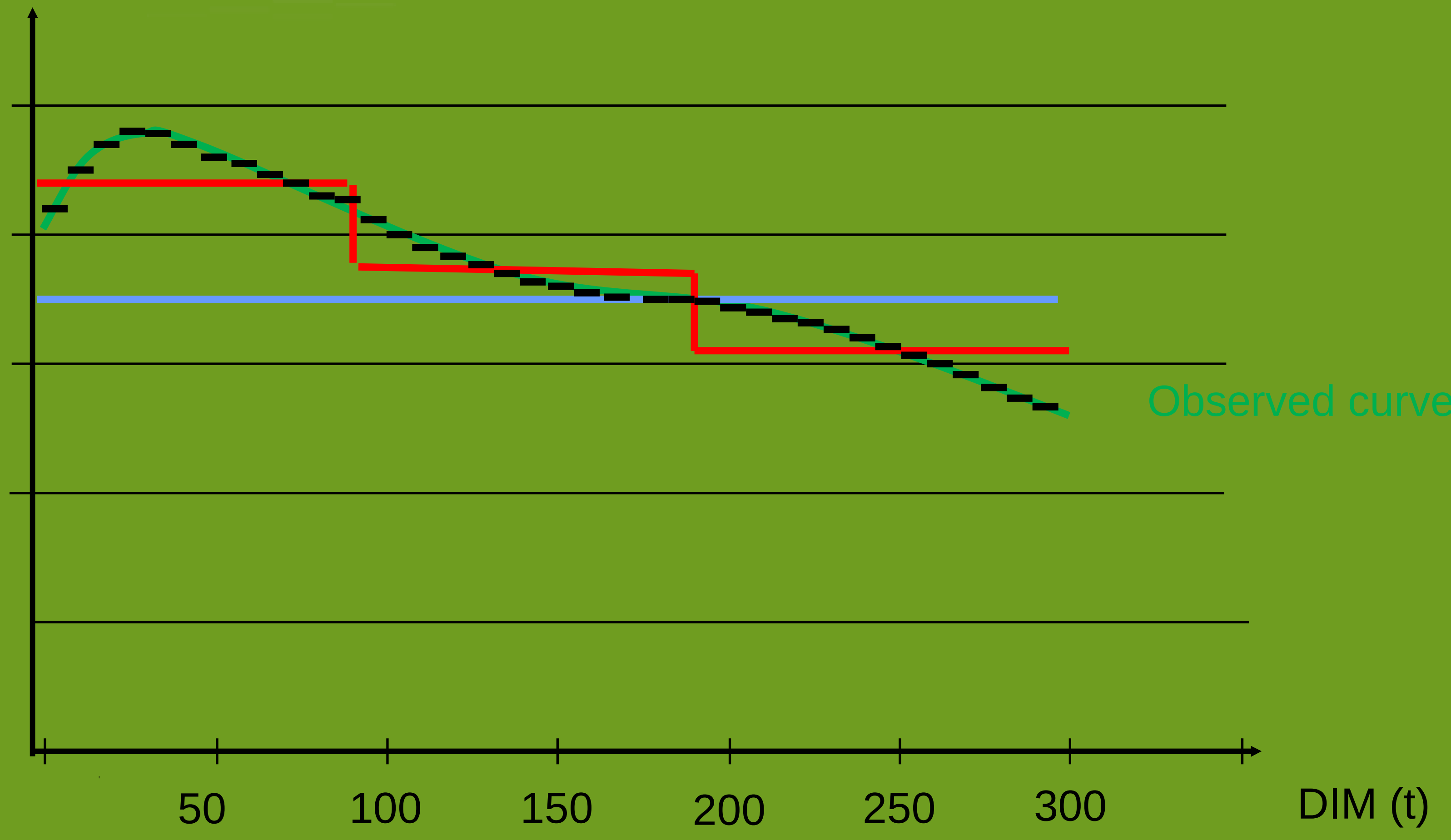
S

see h omputing-Lactation-Yield.pdf

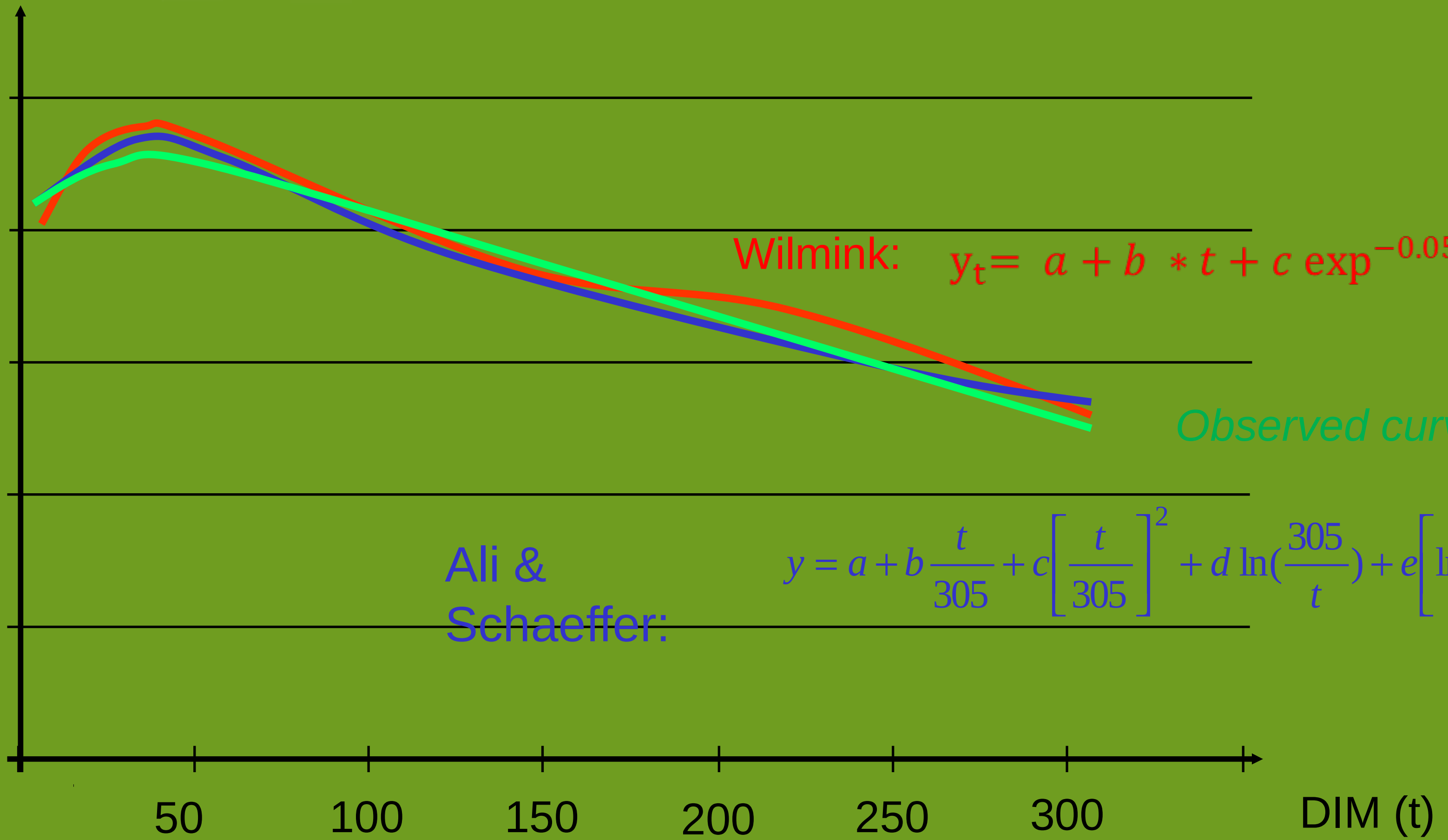
- Test Interval Method (TIM) or “Fletcher method” in the early 1900s (although their reference is Sargent (1968)).

How to compute a *standardized* lactation record?

Modelling lactation curves:
Non parametric curves



Parametric curves



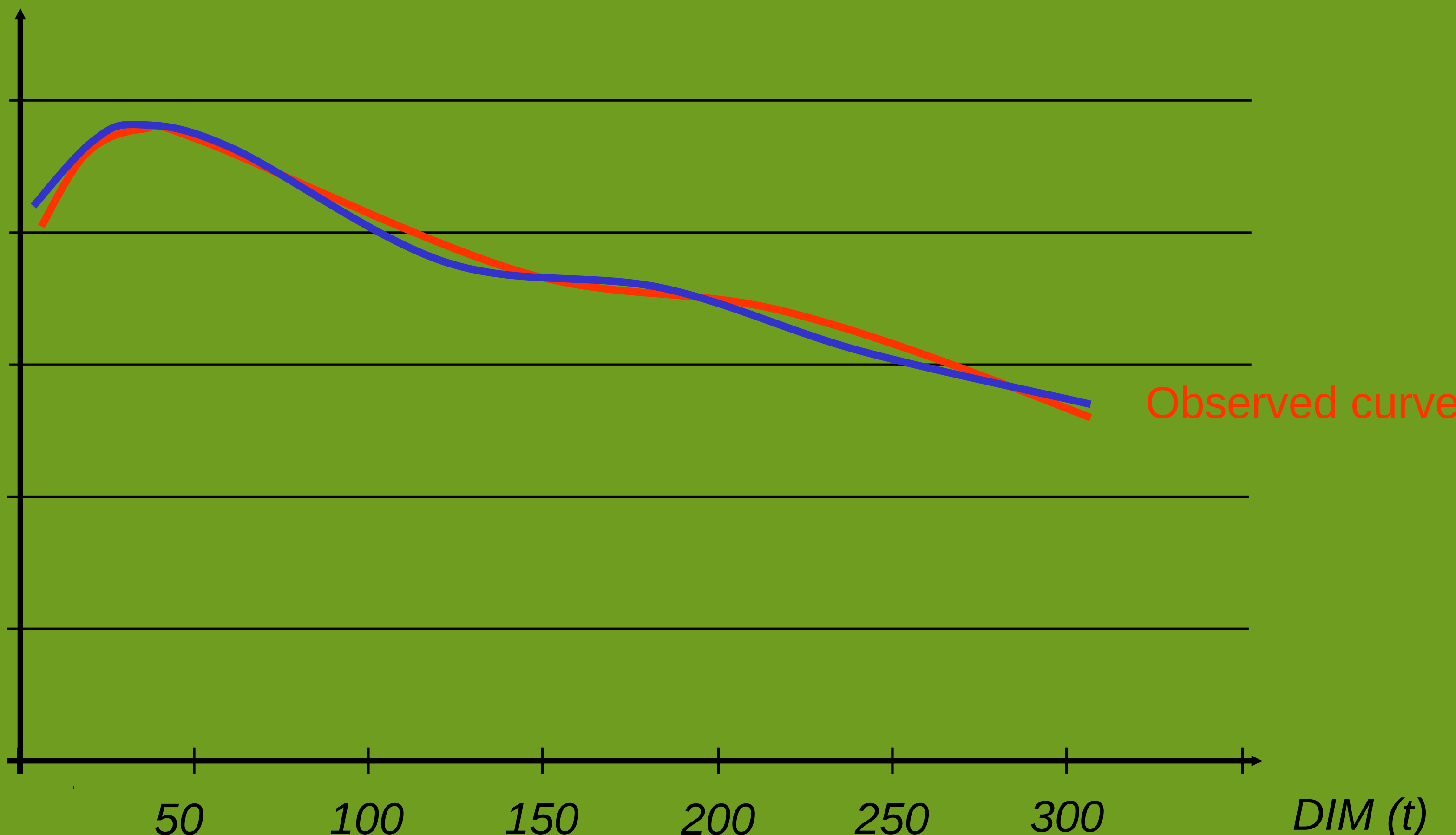
Wilmink: $y_t = a + b * t + c \exp^{-0.05\,t}$

Observed curve

Ali & Schaeffer: $$y = a + b\frac{t}{305} + c\left[\frac{t}{305}\right]^2 + d\ln\left(\frac{305}{t}\right) + e\left[\ln\left(\frac{305}{t}\right)\right]^2$$

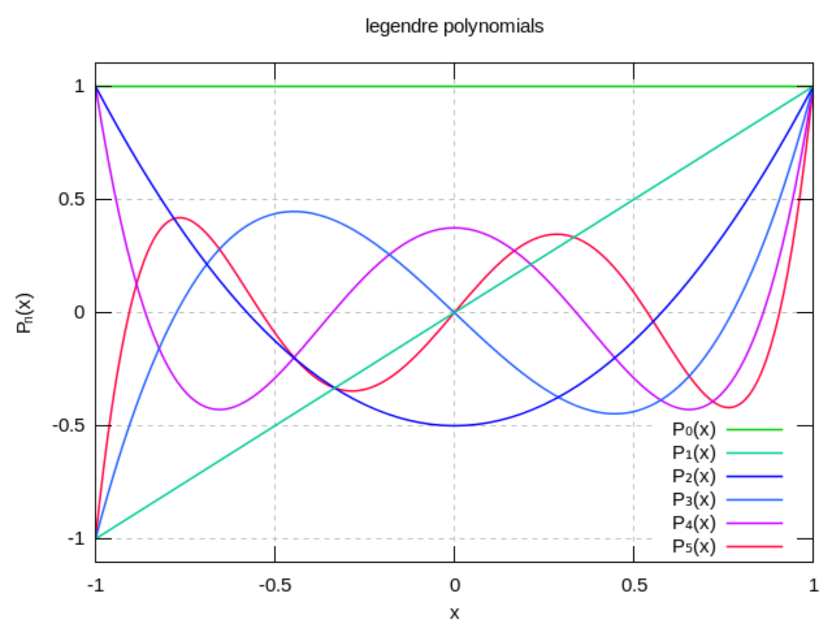
Polynomials

$$y = a + bt + ct^2 + dt^3 + \ldots$$



Legendre Polynomials

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}\left(3x^2-1\right)$
3	$\frac{1}{2}\left(5x^3-3x\right)$
4	$\frac{1}{8}\left(35x^4-30x^2+3\right)$
5	$\frac{1}{8}\left(63x^5-70x^3+15x\right)$
6	$\frac{1}{16}\left(231x^6-315x^4+105x^2-5\right)$
7	$\frac{1}{16}\left(429x^7-693x^5+315x^3-35x\right)$
8	$\frac{1}{128}\left(6435x^8-12012x^6+6930x^4-1260x^2+35\right)$
9	$\frac{1}{128}\left(12155x^9-25740x^7+18018x^5-4620x^3+315x\right)$
10	$\frac{1}{256}\left(46189x^{10}-109395x^8+90090x^6-30030x^4+3465x^2-63\right)$

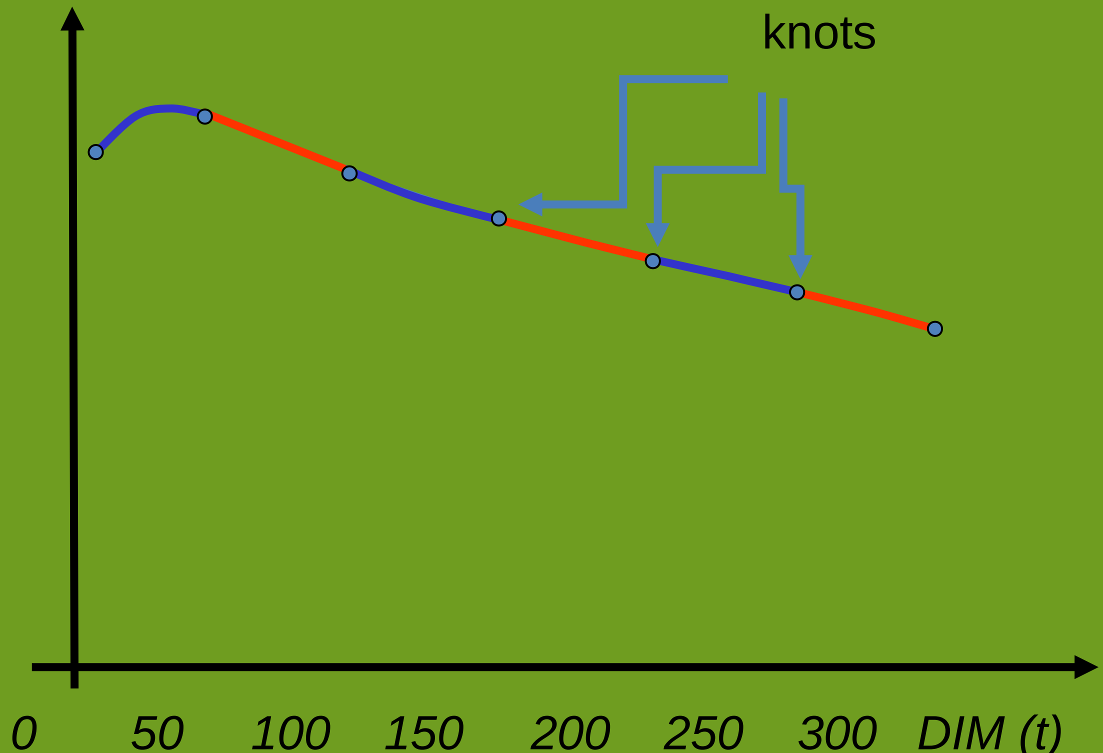


https://en.wikipedia.org/wiki/Legendre_polynomials

Splines

https://en.wikipedia.org/wiki/Spline_(mathematics)

Different types: linear, quadratic, cubic, natural, …



Options for analysis

- **Lactation records**:
  - Actual lactation records or 305d lactation records
  - often based on +/- 10 monthly test dates
  - = trait used in the genetic evaluation
- **Test day records**:
  - Direct analysis of production at each test date
  - No loss of accuracy when extrapolating a lactation
  - No need to standardize / correct for lactation length
  - No need to exclude very short lactations less bias
  - Not a problem if missing test records or if larger intervals between tests

Examples of effects in the model

7337 kg

- Herd x year (2010)
- Month of calving (October)
- Age at calving (in months or in class)
- Additive genetic effect
- Permanent environment effect
- Days between calving and successful AI

22.4 kg

- Herd x test date(15/05/2011)
- Month of production (May)*
- Age at test date*
- Genetic effect*
- Permanent environment effect*
- Number of days in gestation
- Stage of lactation

** Other model options exist (see later)*

Benefits of test day models

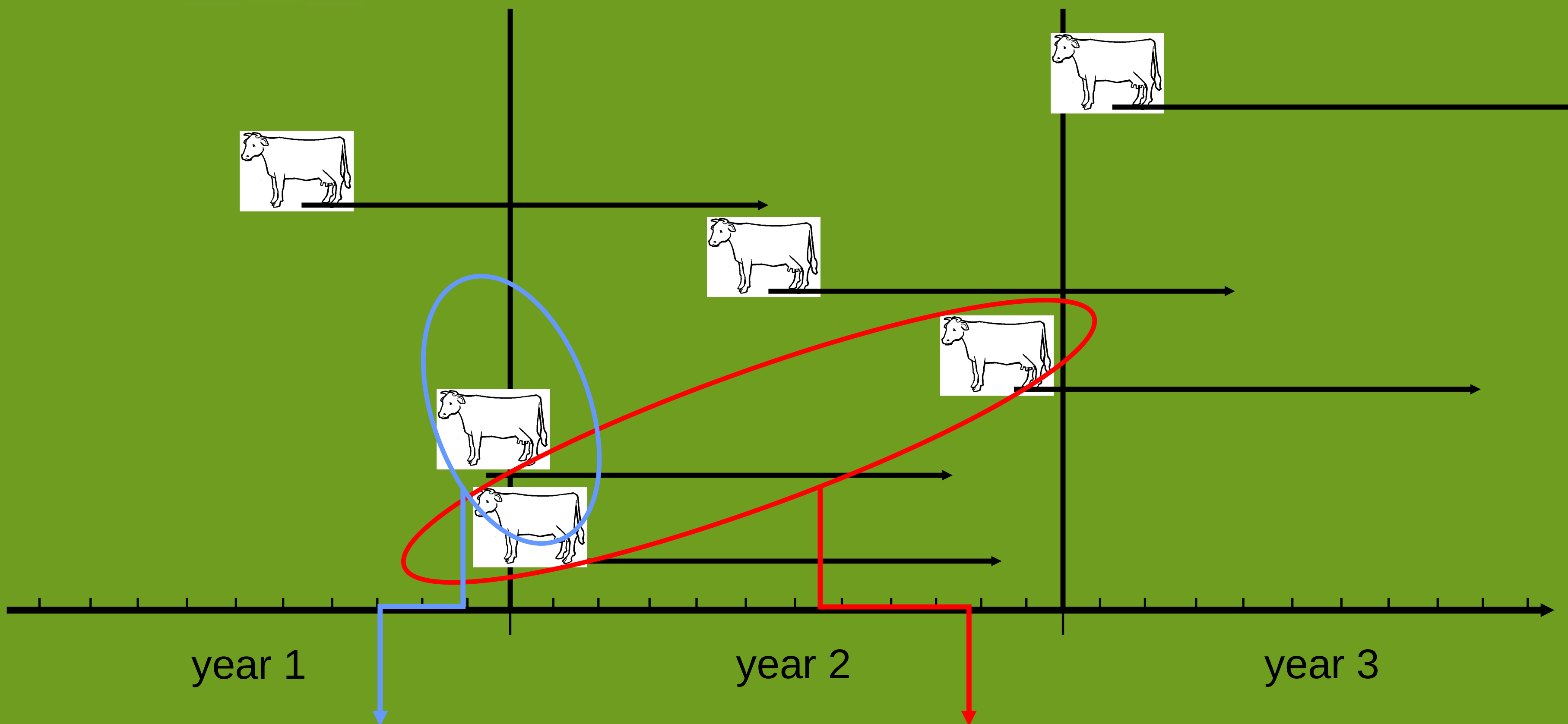
- Each effect influences a particular test on a particular day rather than the whole lactation.
  - e.g.: the weather on a particular day of test influences production of all cows in the same « **contemporary group** »

    and this group changes at each test date
  - the stage of gestation has an effect only when the cow is actually pregnant.
- For each observation (test date), a particular effect can change
  - e.g.: the age class changes by +1 month at each test date.
- An effect can change depending on the stage of lactation
  - e.g.: the influence of the month of production may be different between the lactation peak or at the end of lactation

 All this will reduce the (unexplained) residual variance

Contemporary group definition

= all cows starting in the same **herd** on the same **year**

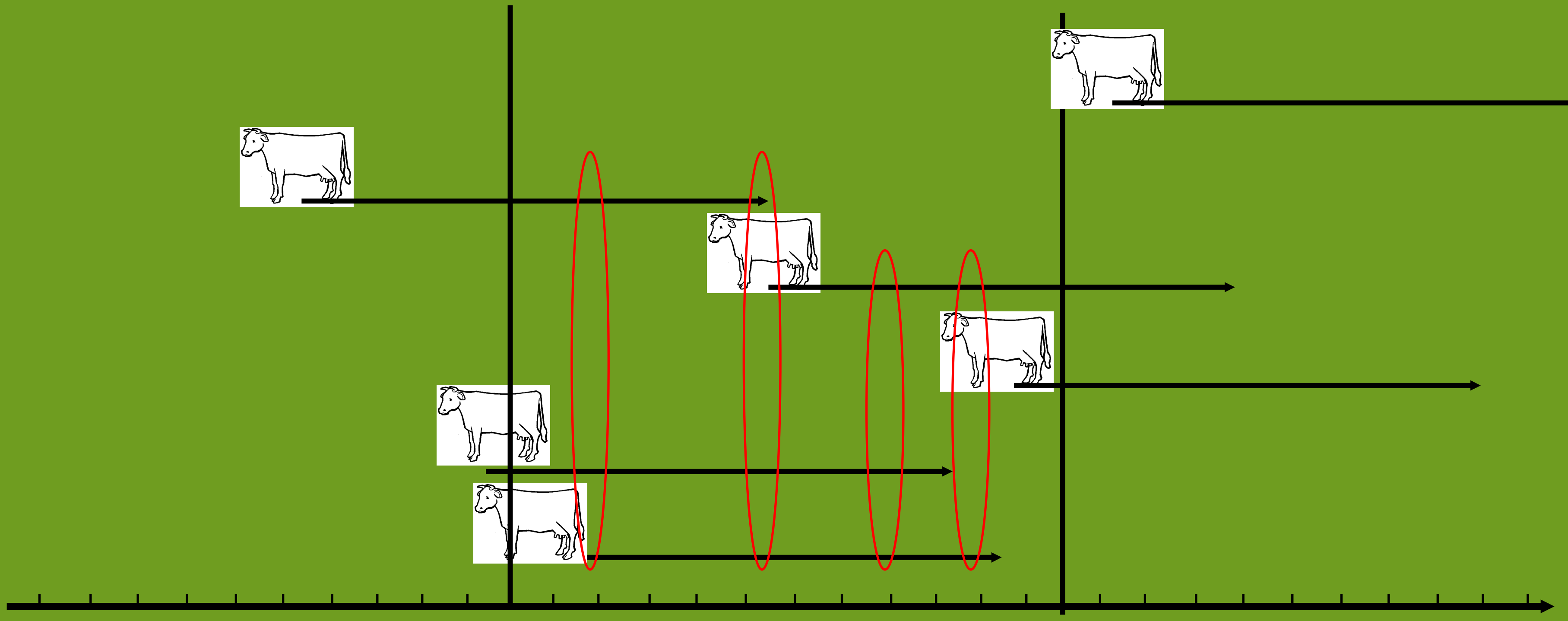


These 2 cows are in the same contemporay grou

These 2 cows are not in the same contemporay group

Contemporary group definition

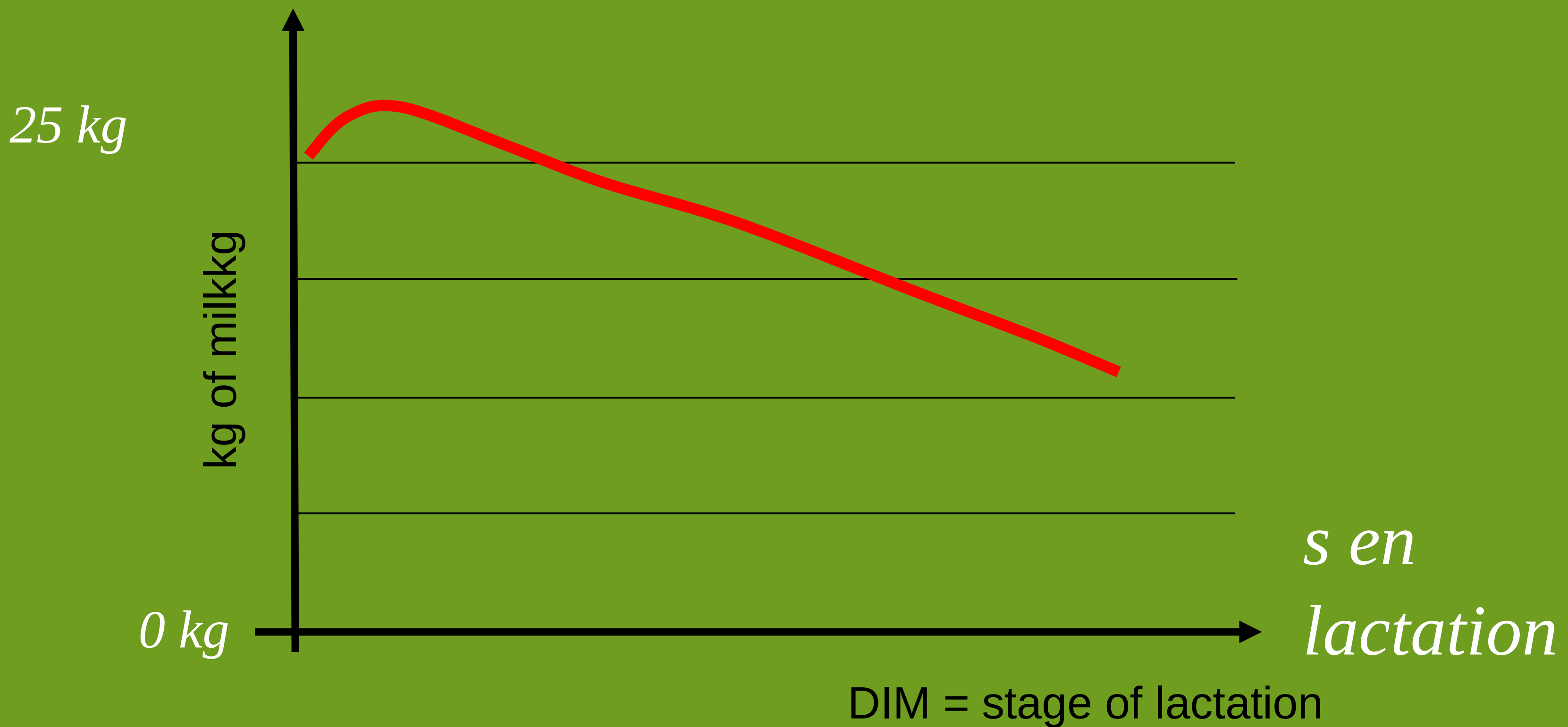
= all cows milked in the same **herd** on the same **test date**



The **herd x test date effect** is more precise than the **herd x year effect**

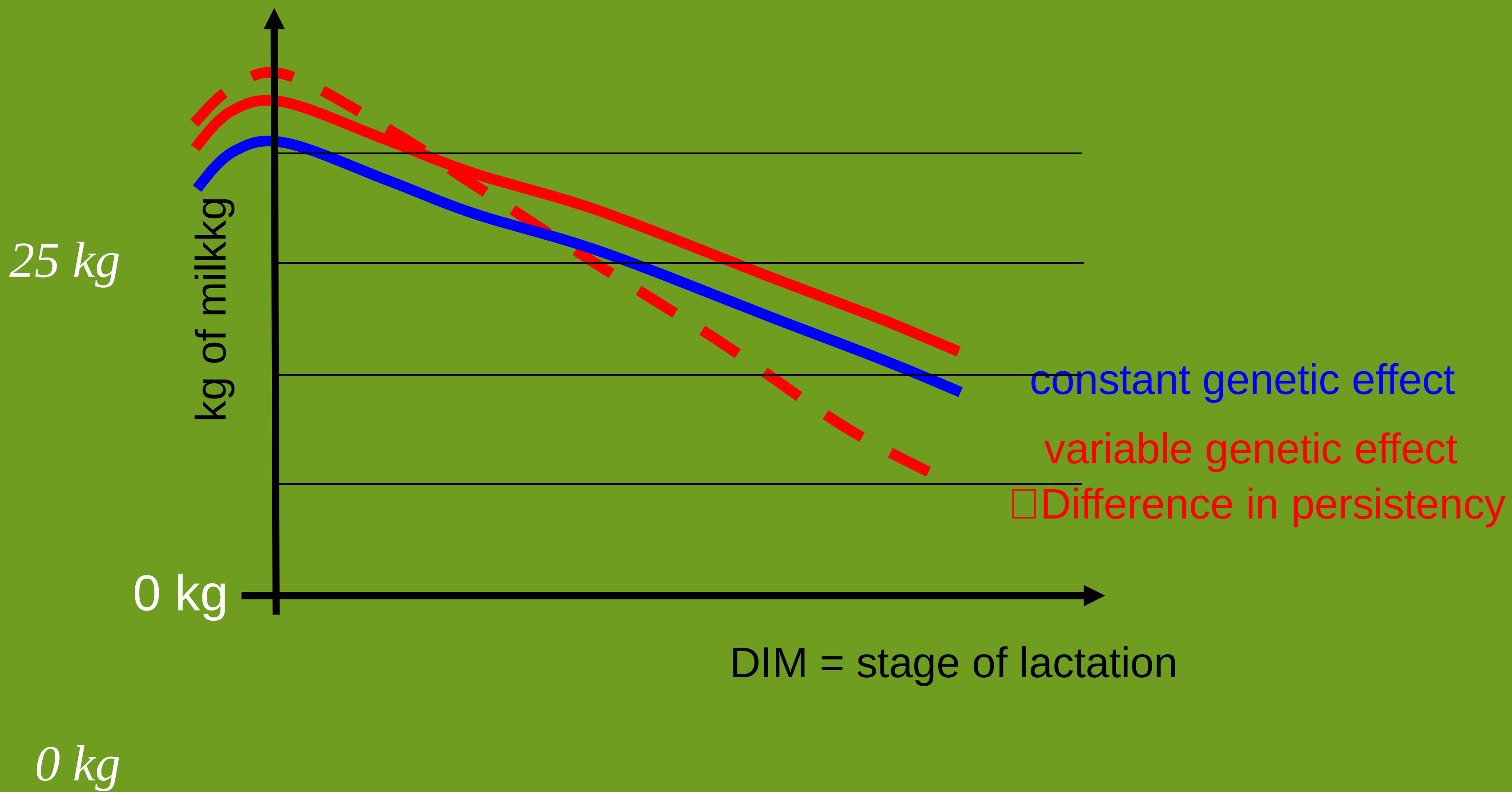
Shape of lactation curve

Base curve: accounts for stage of lactation



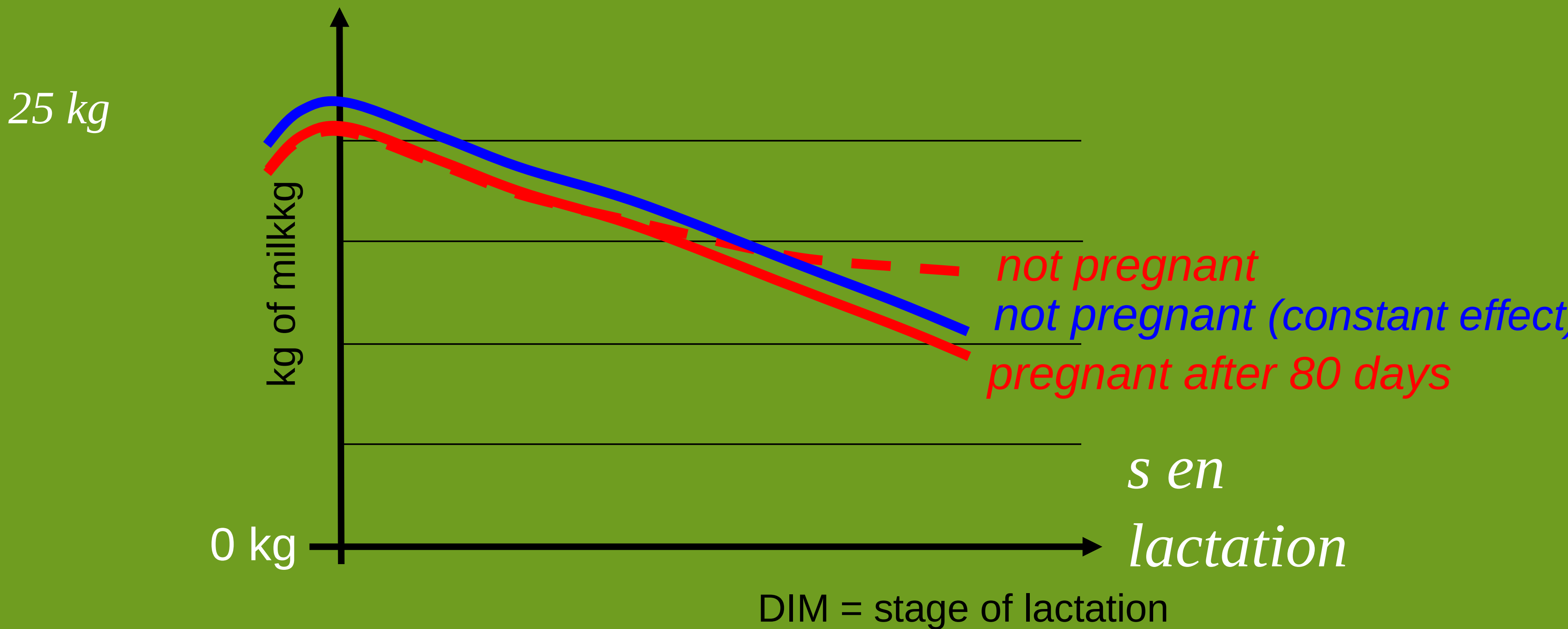
Shape of lactation curve

Genetic effect



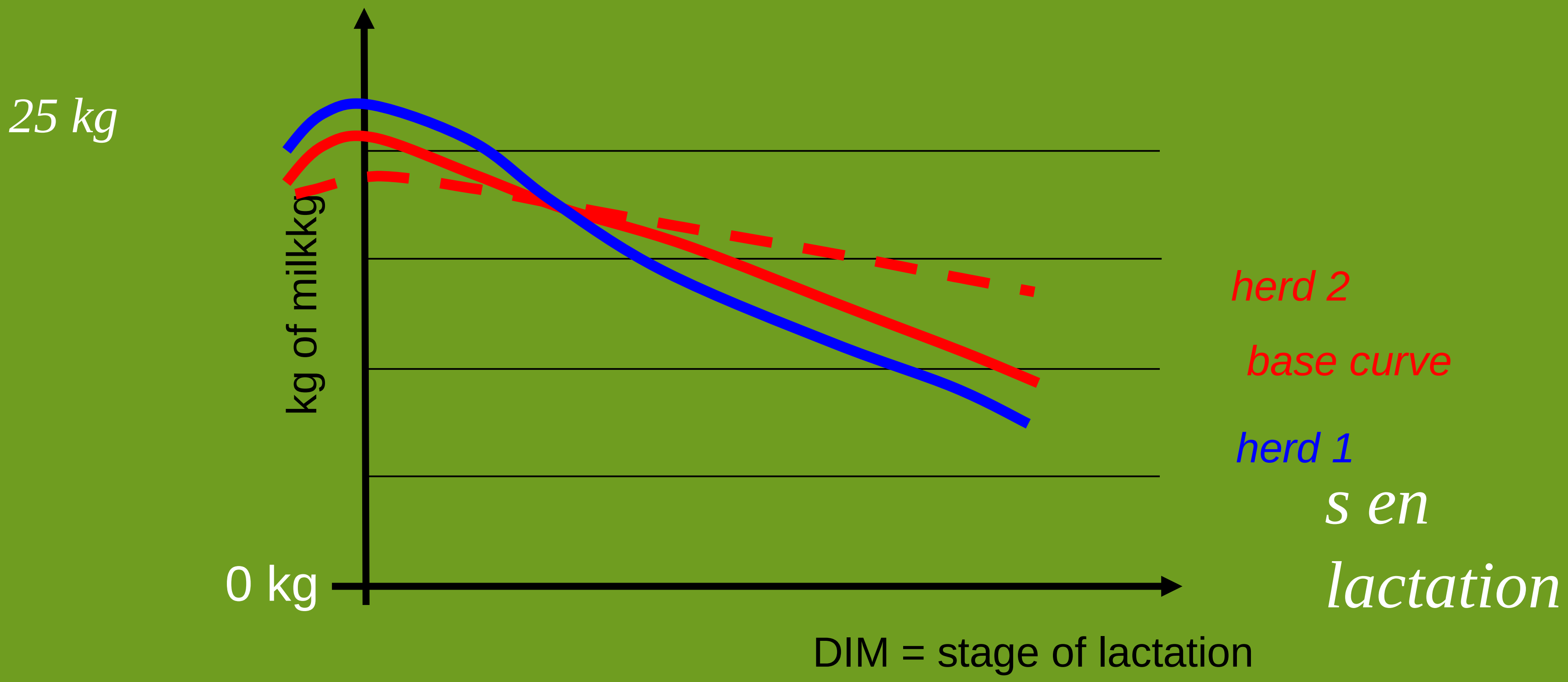
Shape of lactation curve

Gestation effect



Shape of lactation curve

Herd



vincent.ducrocq@inra.fr

Modelling fixed effects

- Examples: herd x test date, lactation number, age on test date or at calving, stage of gestation, length of dry period
- Some effects are assumed constant over the lactation, others vary during lactation
- Options: assume additive effects or interactions

Example of interactions

Age at calving

2 year old

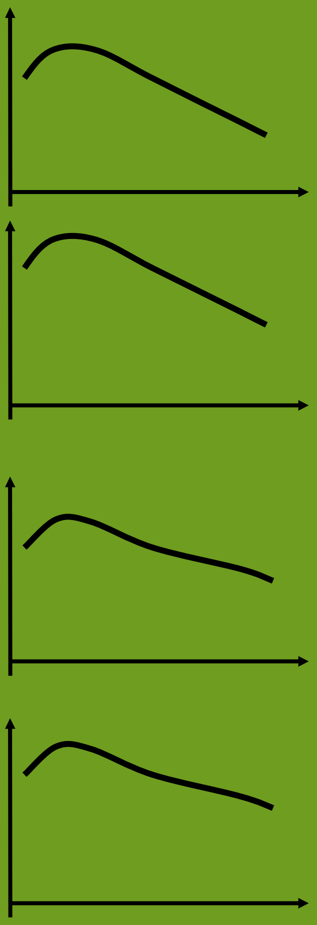
3 year old

x

Season of calving

January-June

July-December



2 year old
x janvier-juin

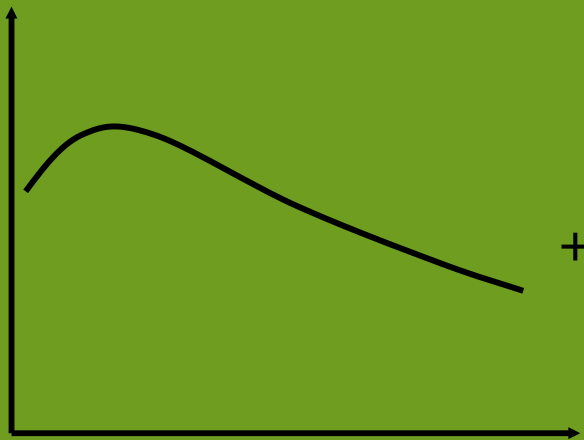
3 year old
x janvier-juin

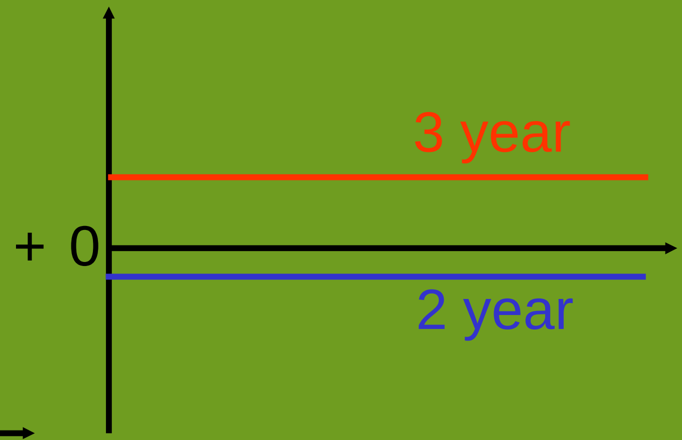
2 year old
x July-December

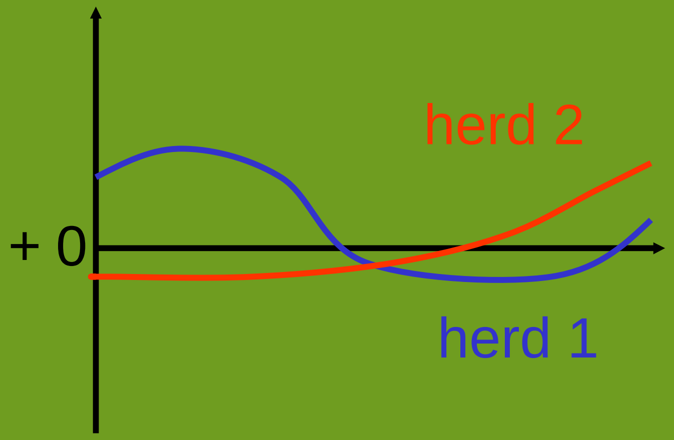
3 year old
x July-December

Example of additivity

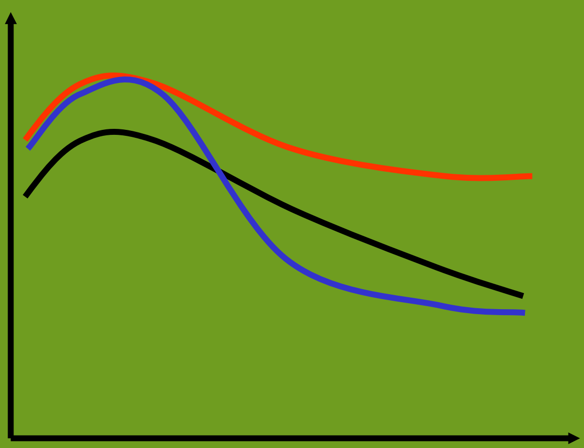
base curve + constant effect + variable effect







=



Curves
for combinations
of effect

Modelling the genetic part

- Constant genetic effect over the lactation : not realistic

= Fixed regression model (only the environmental part is described by curves). **Abandoned almost everywhere …**

- Constant genetic effect over parts of the lactation, for example within each month of lactation but different from one month to the next.

= Multiple trait model (production during each month = 1 distinct trait)

Useful to check results from random regression models. Heavy computing.

- Continuous genetic variation during the lactation

the genetic effect is described as linear function of DIM

= Random regression model

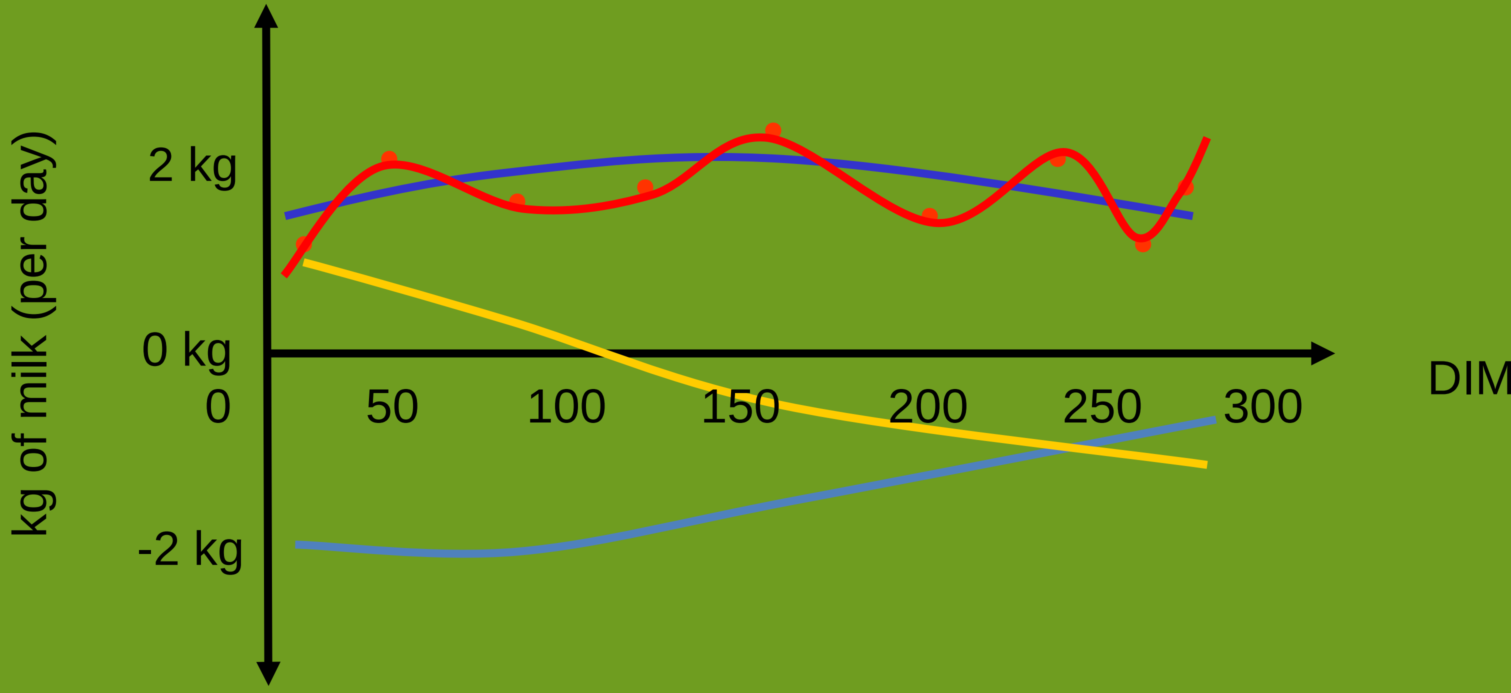
Random regression model

- Varying genetic effect :
  - Change in persistency
  - A genetic value per day, per month?
  - Summarize the information in a curve, with a limited number of parameters
  - possibility to summarize the curve to 2 parameters (an average over the lactation + a measure of persistency)
- Persistency : how to measure it?
  - difference between two dates (e.g., day 60 and 280)
  - Production ratio (independent from level of production

    e.g., prod 1-100 days / prod 101-200)
  - if simple curve: juste the slope of the curve

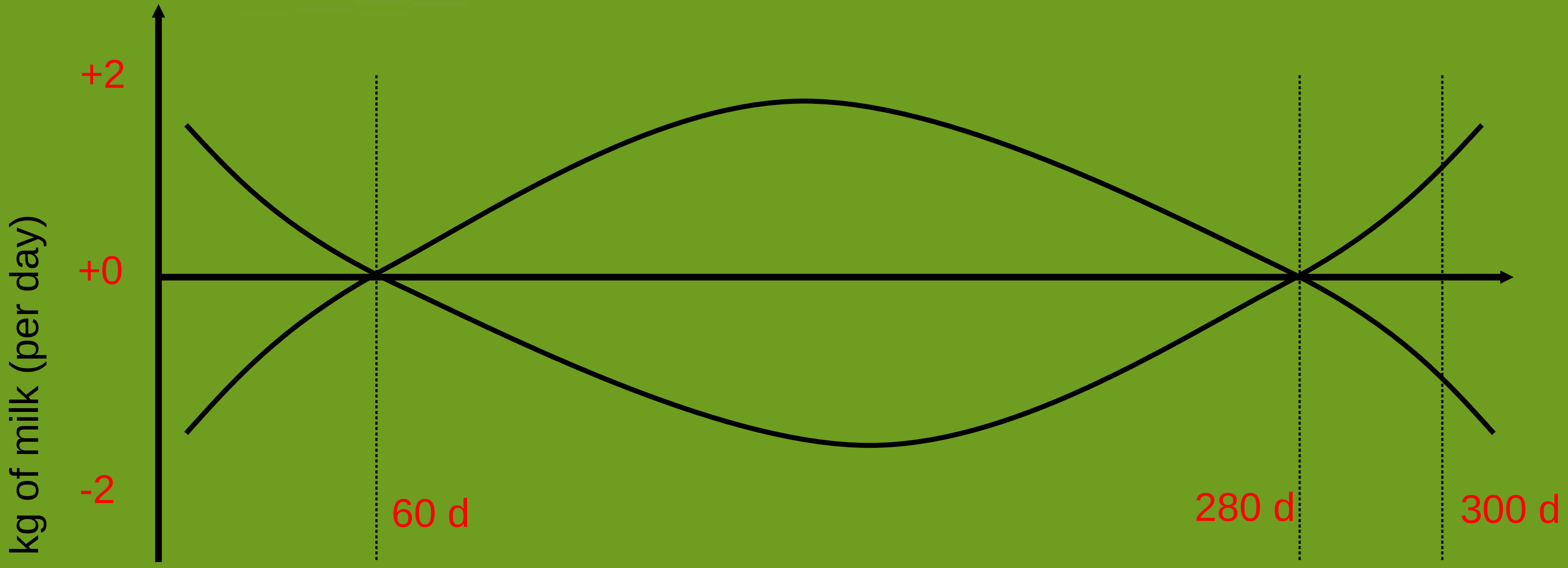
Random regression models

- Varying genetic effects :
  - A genetic value per day, per week, per month?
  - Summarize the information into a curve, with a limited number of parameters
  - Accounts for change in persistency



How to measure persistency?

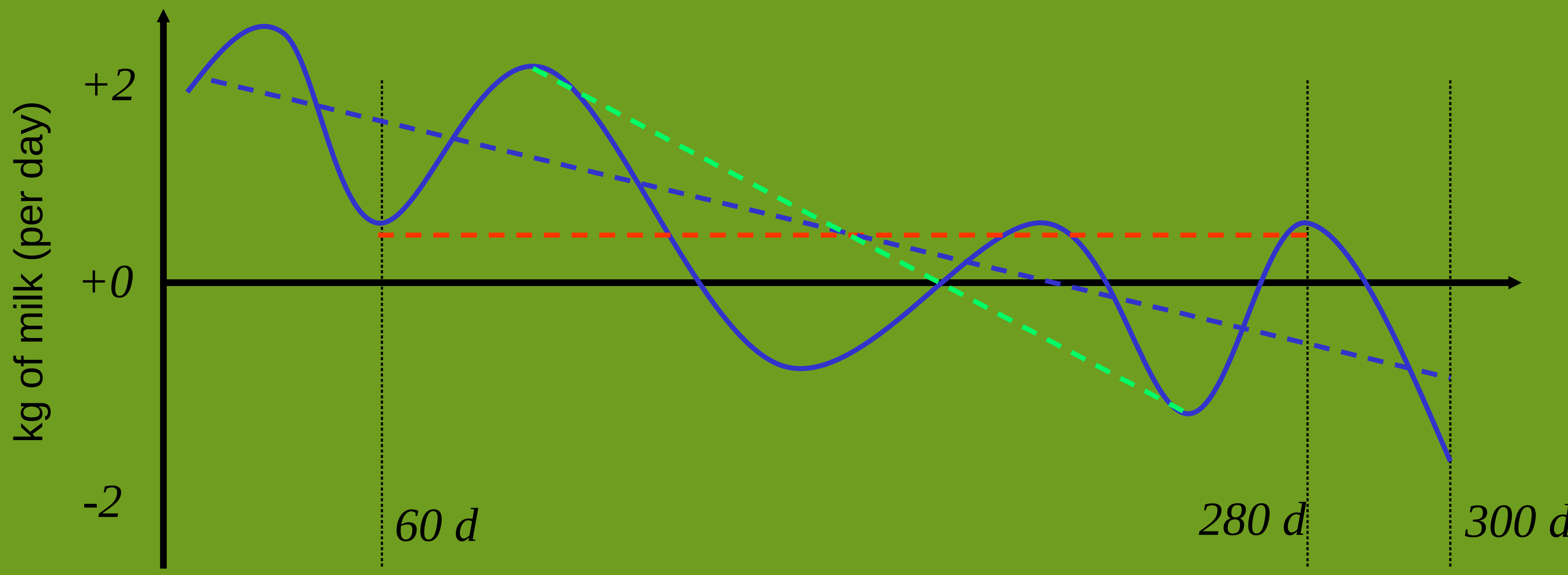
- e.g., production on day 60 / production 280



- = 0 for these three animals…

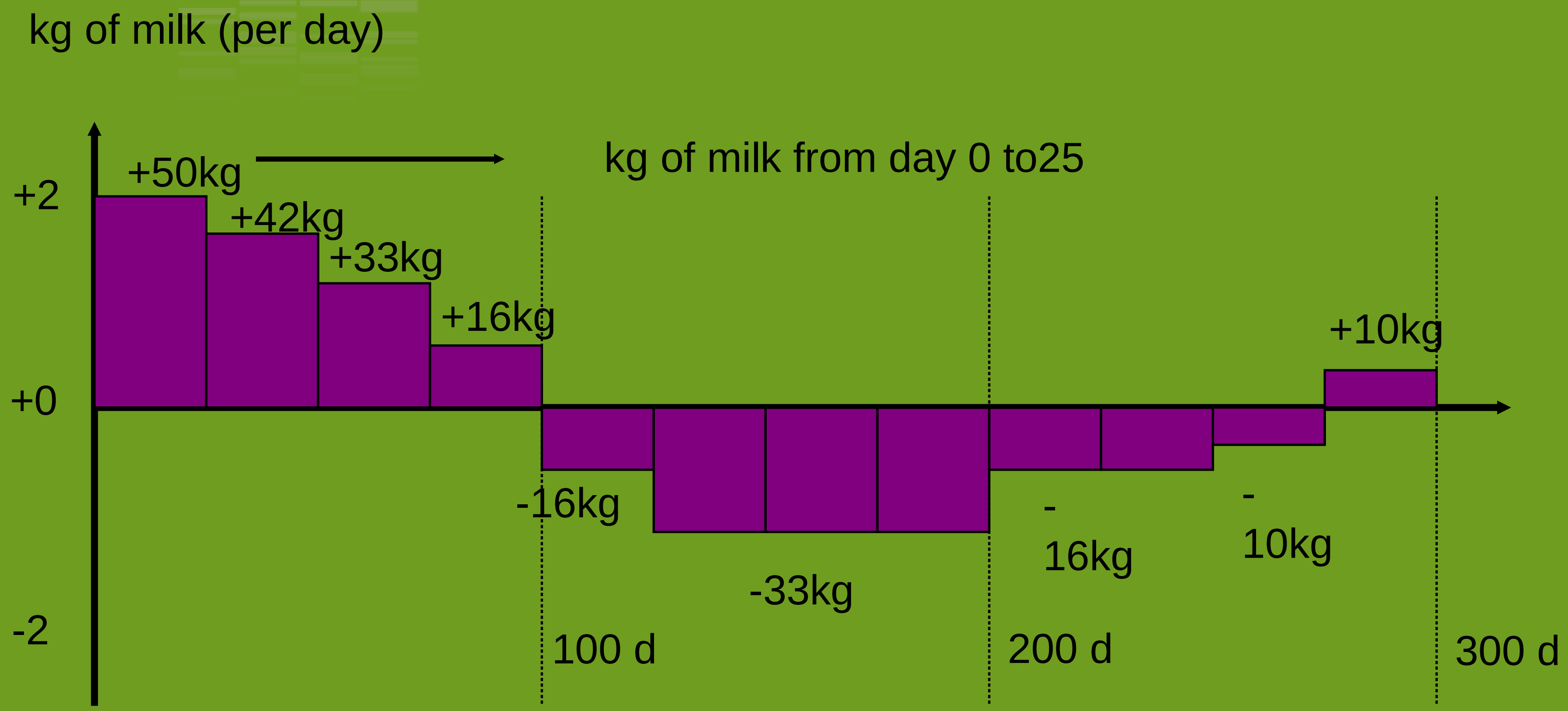
How to measure persistency?

- e.g., slope between two dates?



- From very good to very bad for this animal, depending on the dates

Example of breeding values

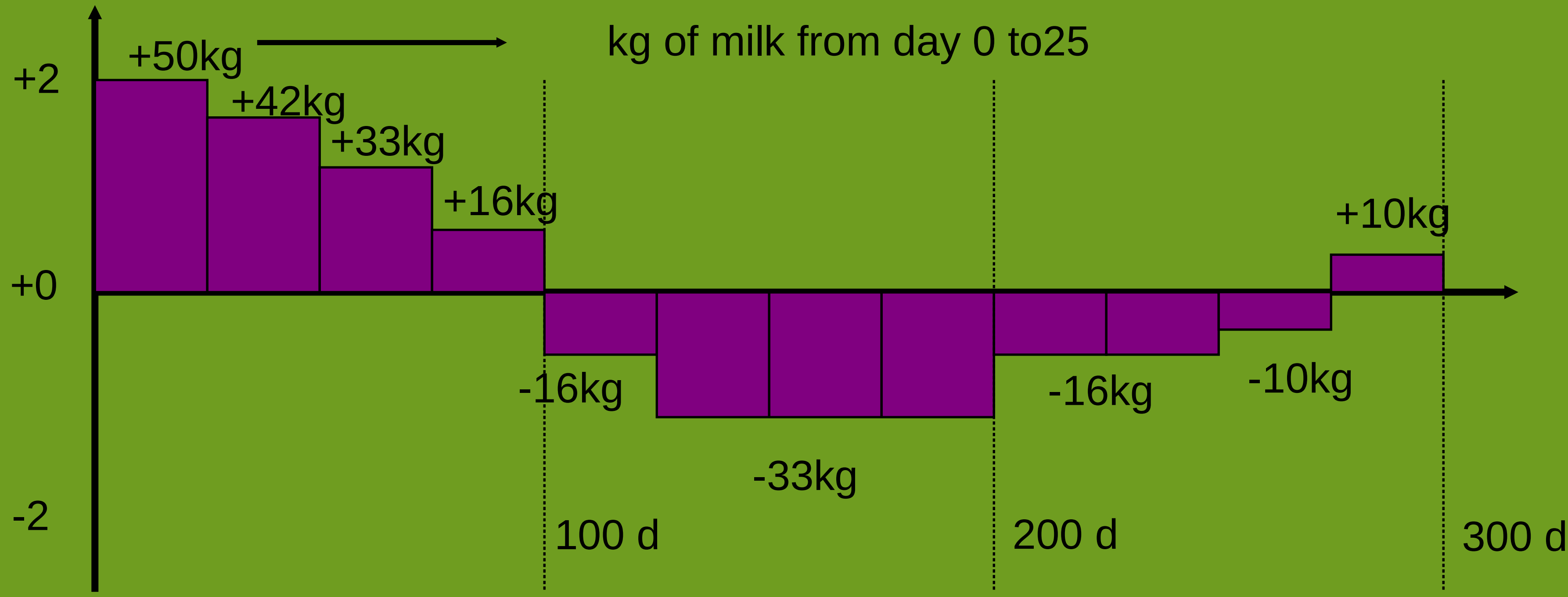


P305 = -6 kg

P100-200 = -115 kg

Example of breeding values

kg of milk (per day)



Persistency 1: diff 105-205 = 0, diff 95-195 = 0.64-(-1.32)= -1.96 kg /day

Persistency 2= $\frac{P200-300+\bar{f}3}{P0-100+\bar{f}1}-\frac{\bar{f}3}{\bar{f}1}$

$\bar{f}1$ is the average prod for the whole population from 0 and 100d, $\bar{f}3$ from 200 à 300 d

Genetic parameters

- Genetic, permanent environment and residual variances (and therefore heritability) change over the lactation and between lactations!
- Also the correlations between genetic effects vary between different periods of the lactation
- Assume a (Legendre) polynomial with 5 parameters

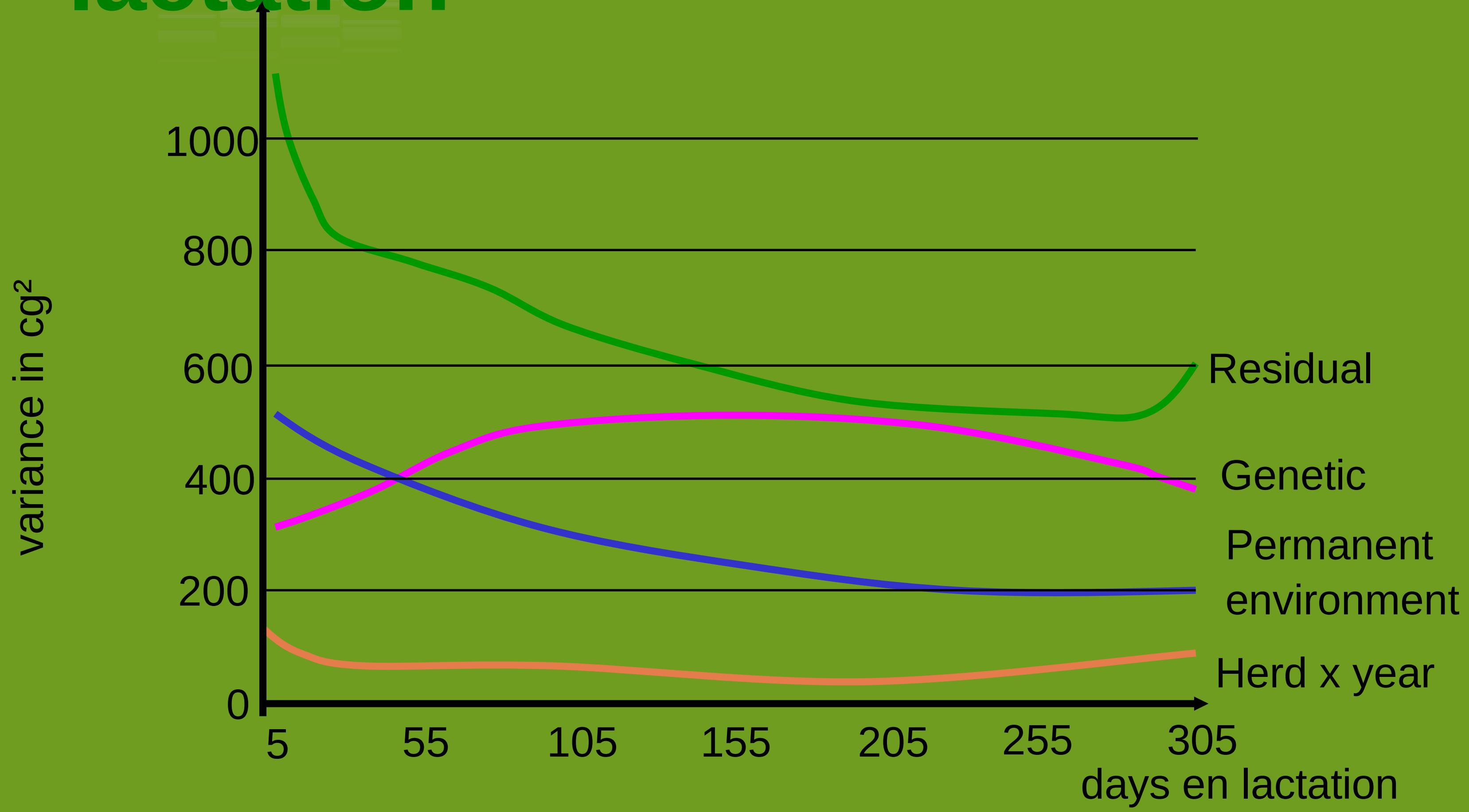
  for the genetic and permanent environment part

  + one residual variance / month of lactation
- 3 lactations => 2 x [(5 x 4) /2] +10 = 30 (co)variance components !

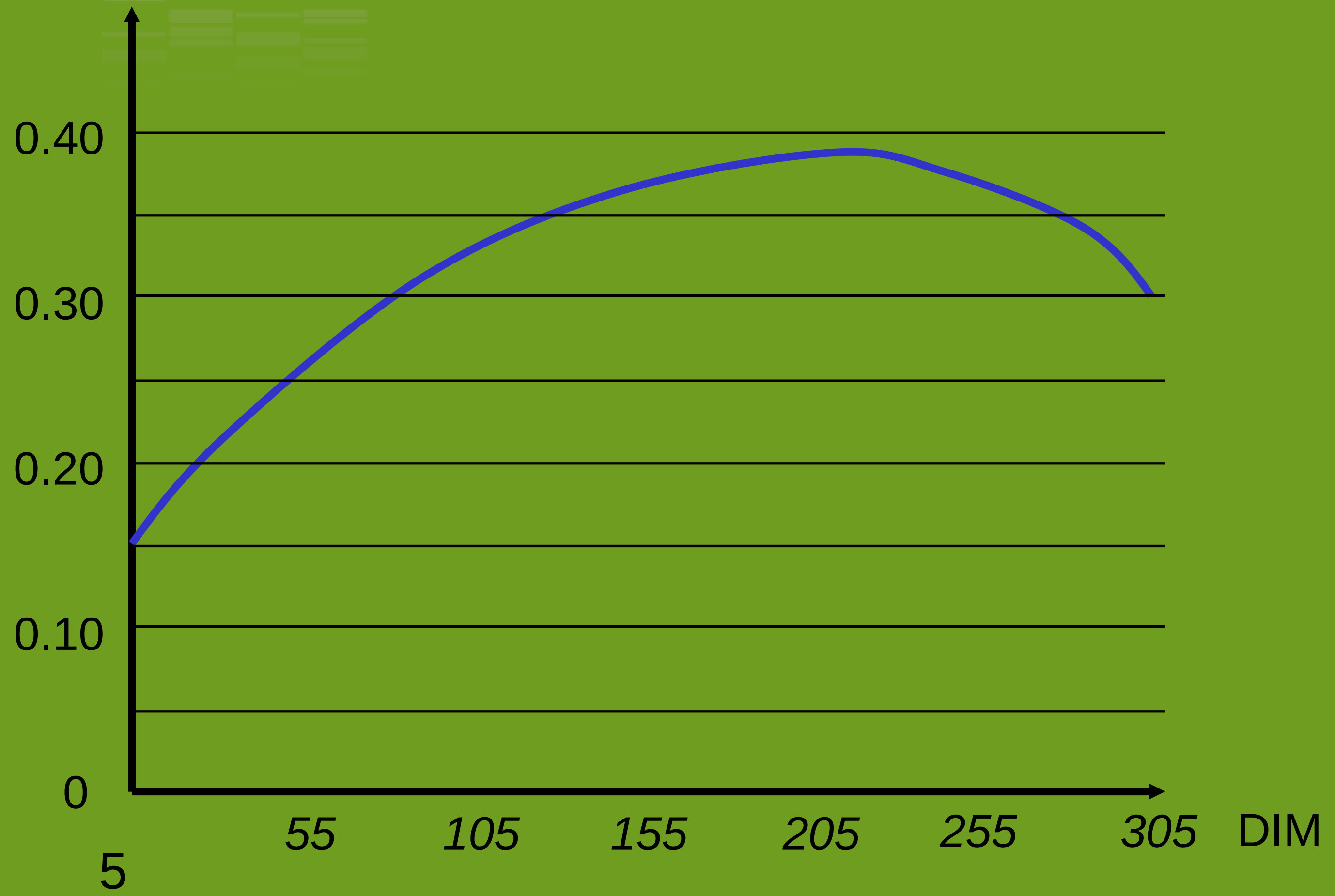
 Large computing times and memory requirements
- Data sets sometimes too small and/or non representatives

heritabilities and correlations possibly strange

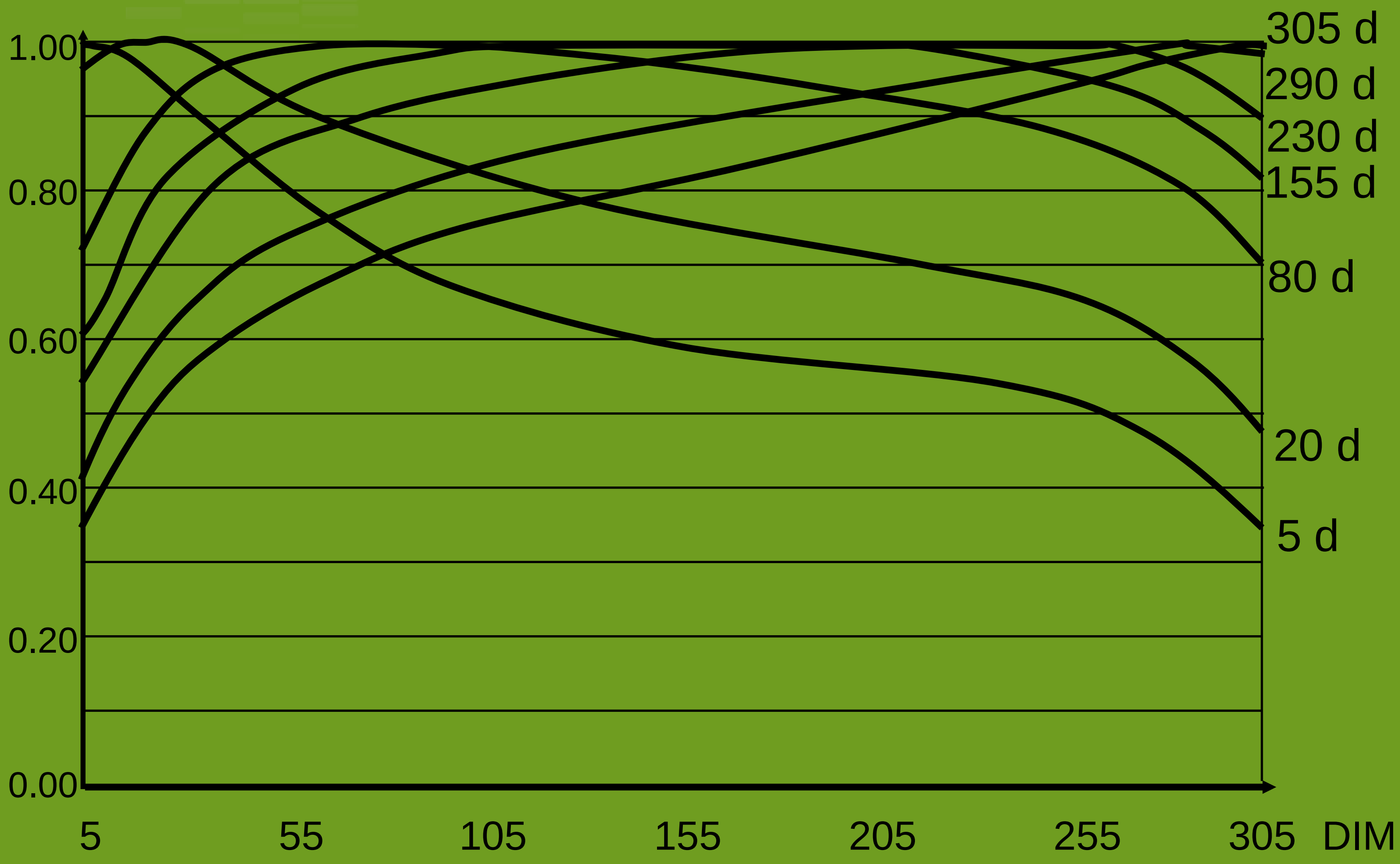
Example: variances in first lactation



Heritability in first lactation



Correlations in first lactation



In pratice…

Production for one test day =

FIXED

- Herd x Test day (HTD)
- Curves region x lact x effects*
- Constant region x year x lact x effects

(effects* =age at calving,
month of calving,
length of previous
dry period (for lact 2, 3)
stage of gestation)

RAND.

- Genetic effect
- permanent environment effect
- Herd x year effect

RAND.

- Genetic effect
- permanent environment effect
- Herd x year effect

Data selection

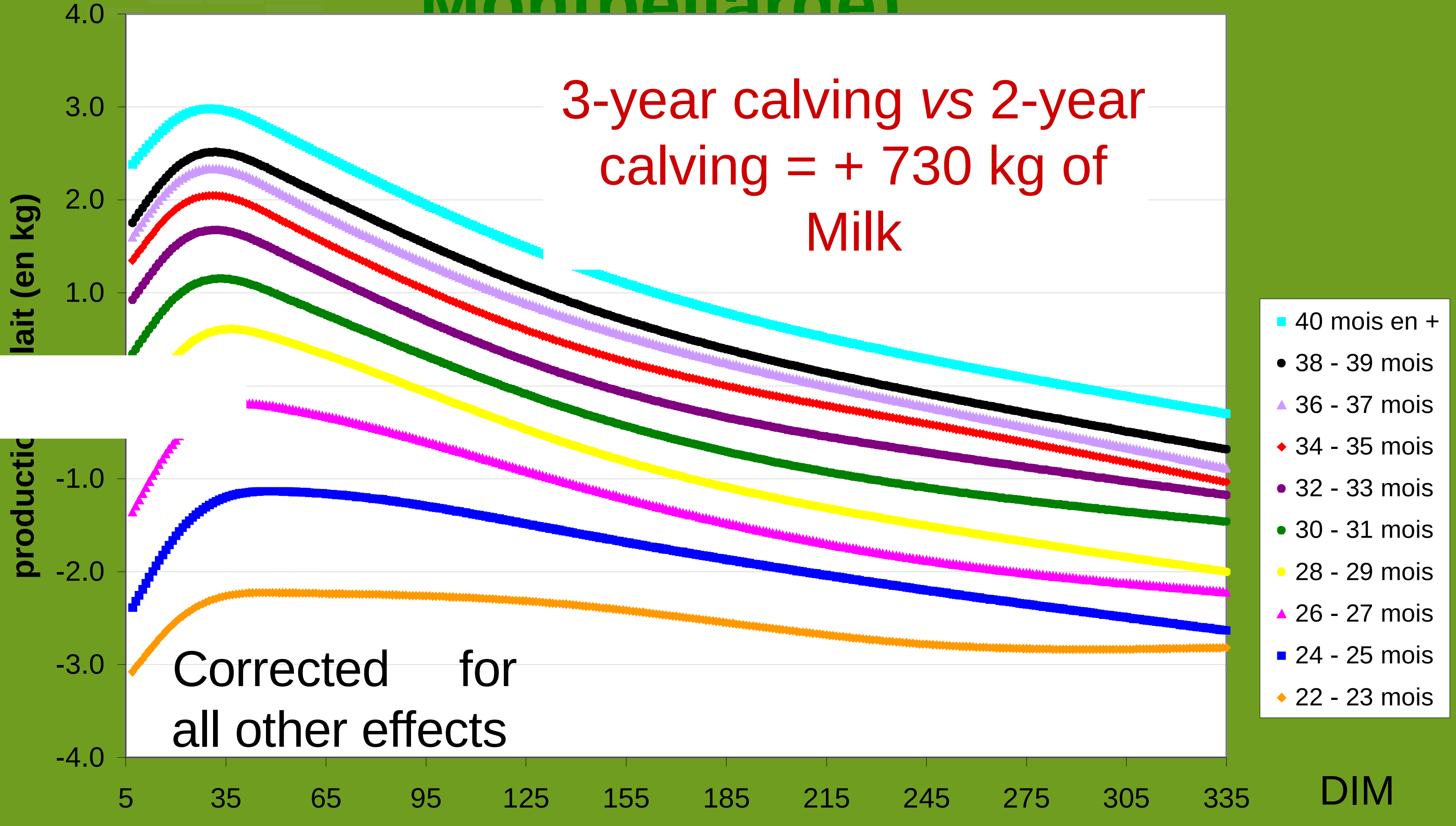


	Montbéliarde (in millions)	Normande (in millions)	Holstein (in millions)
Lactations	6,63	5,71	37,81
TD (initial)	39,05	37,25	256,86
TD (L1-L3 selection 1)	27,33	24,88	170,55
TD (L1-L3 selection 2)	26,17	23,44	167,47
Nb evaluated animals	1,77	1,76	11,86

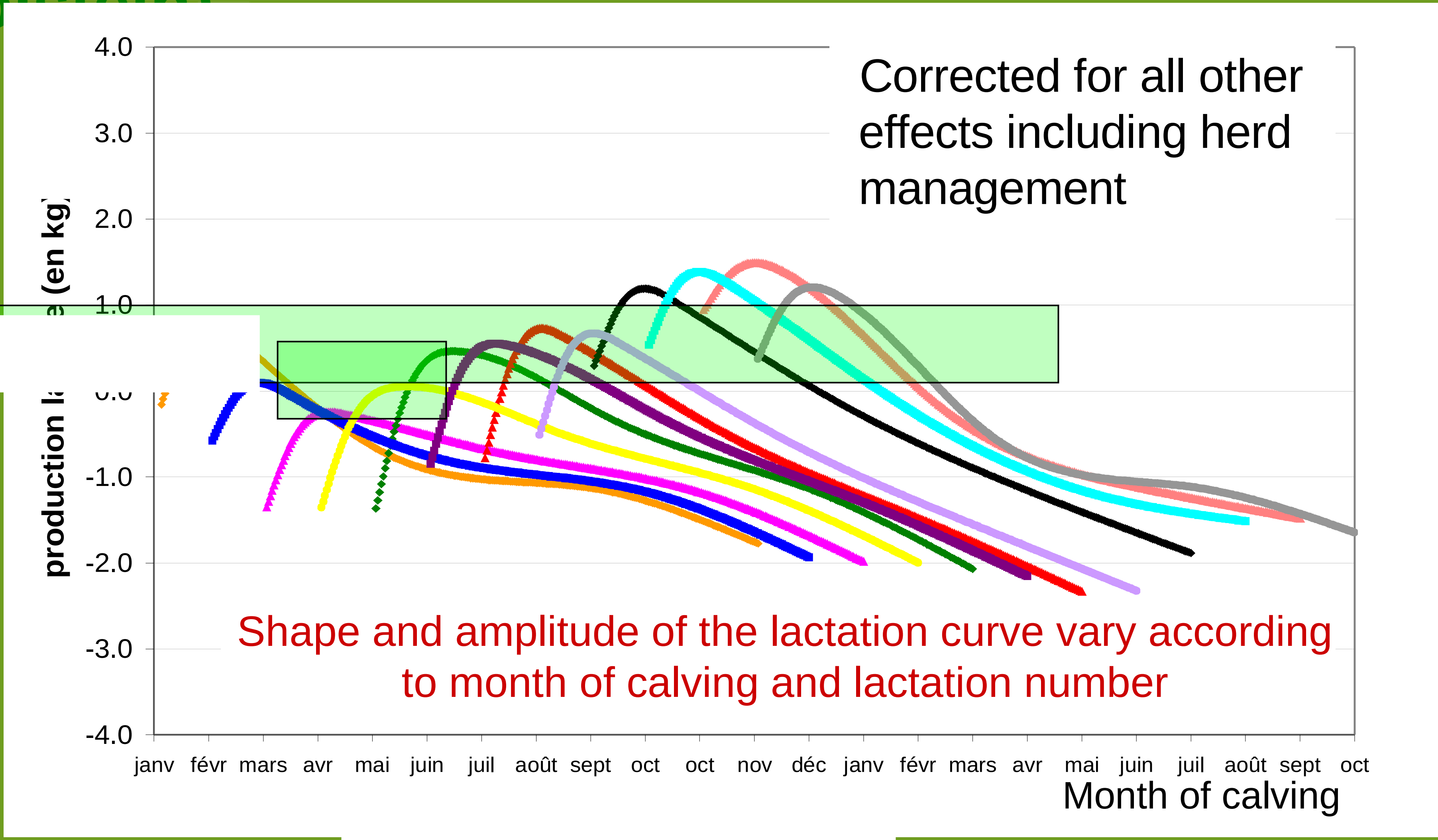
Selection 1 = cows born after January 1 1988, known sire,
stage of lactation between 7 and 335 days, known birth date,
recorded in L1,

Selection 2 (iterative) minimum number= 3 TD per cow (envt perm),
3 TD per test day (HTD), 30 TD per herd x year (HY),
250 TD per class of lactation curve, 100 TD per year class cst

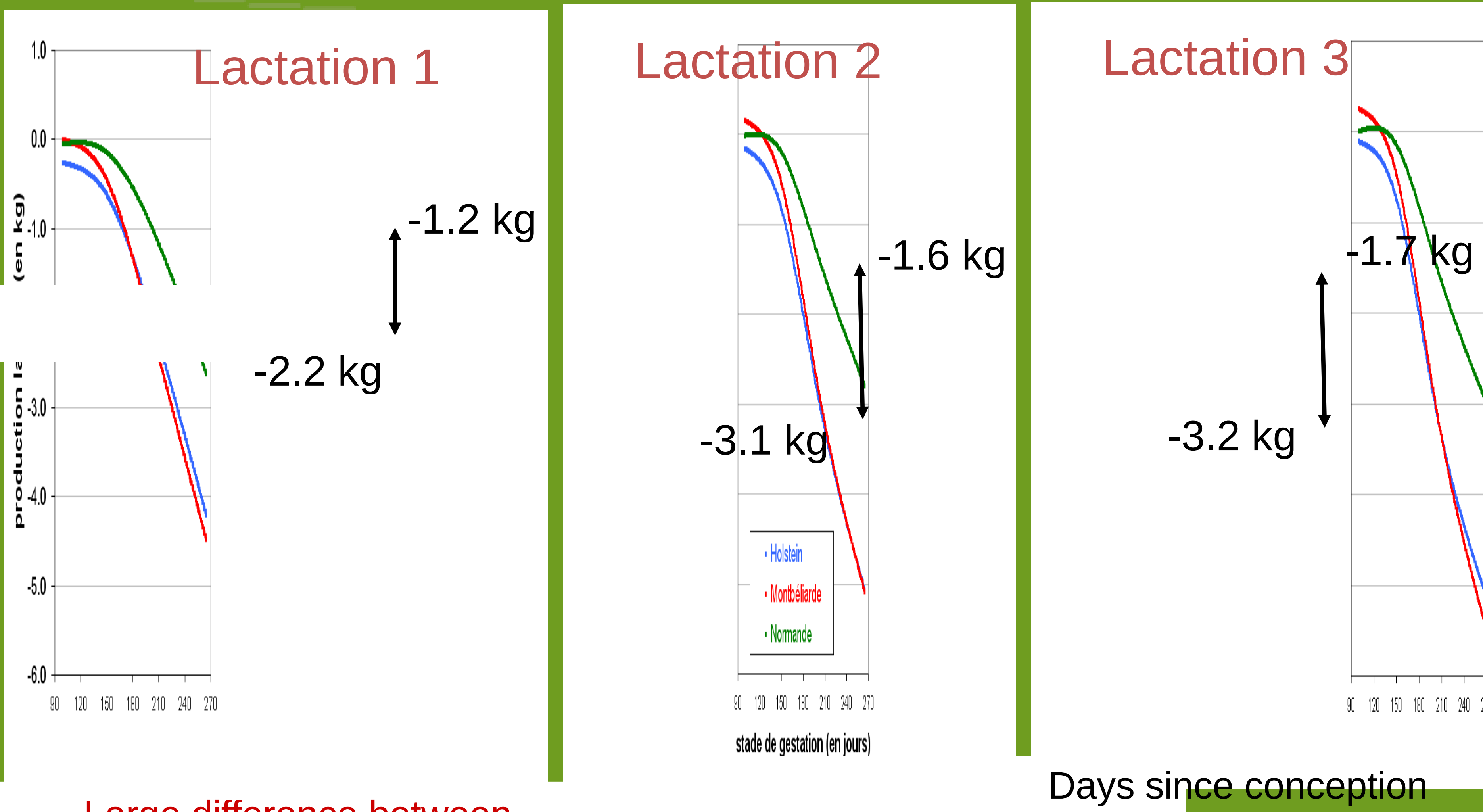
Age at first calving (Lact 1 - Milk - Monthéliarde)



Month of calving effect (Lact 1 -Milk - Holstein)

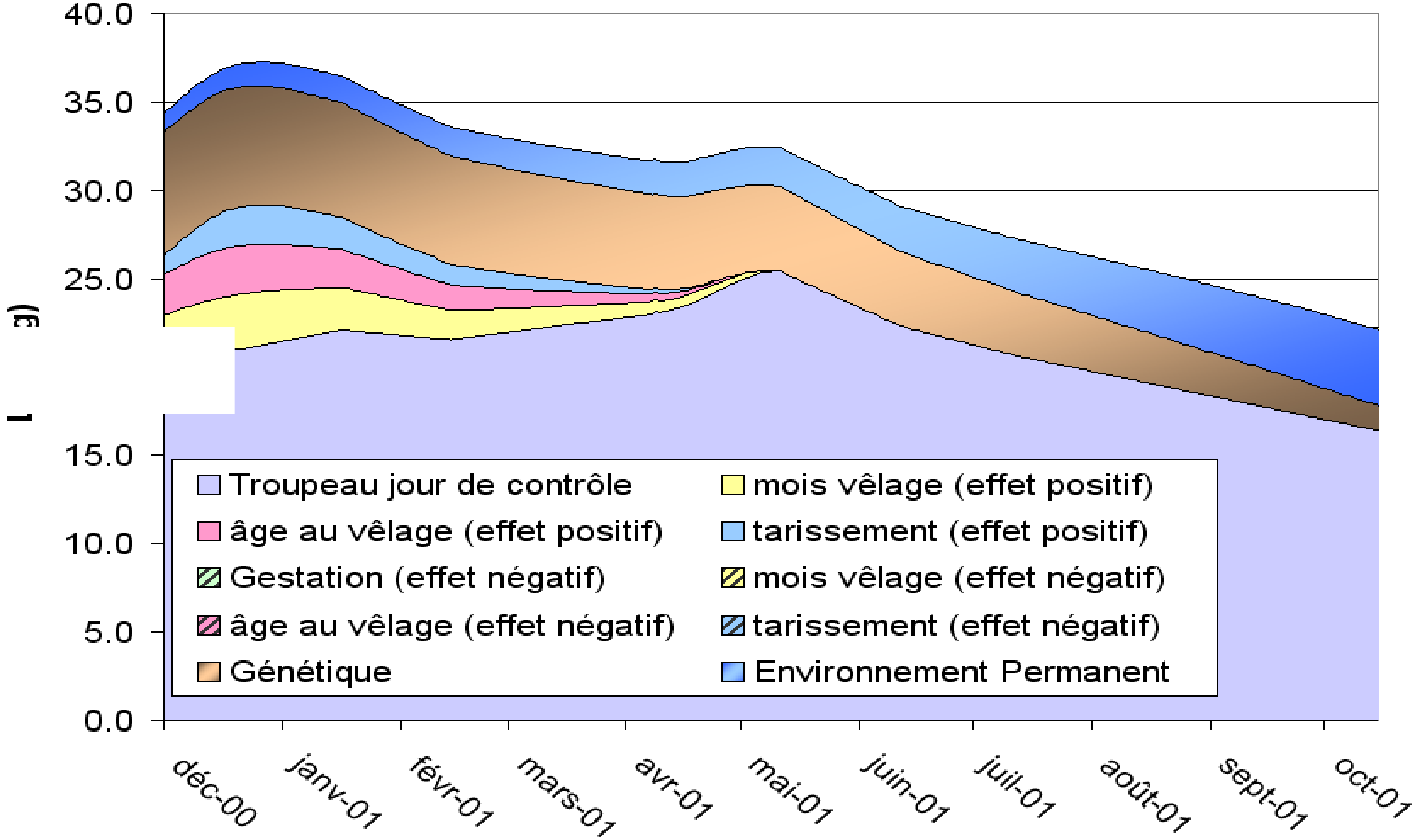


Effect of gestation length (Milk)

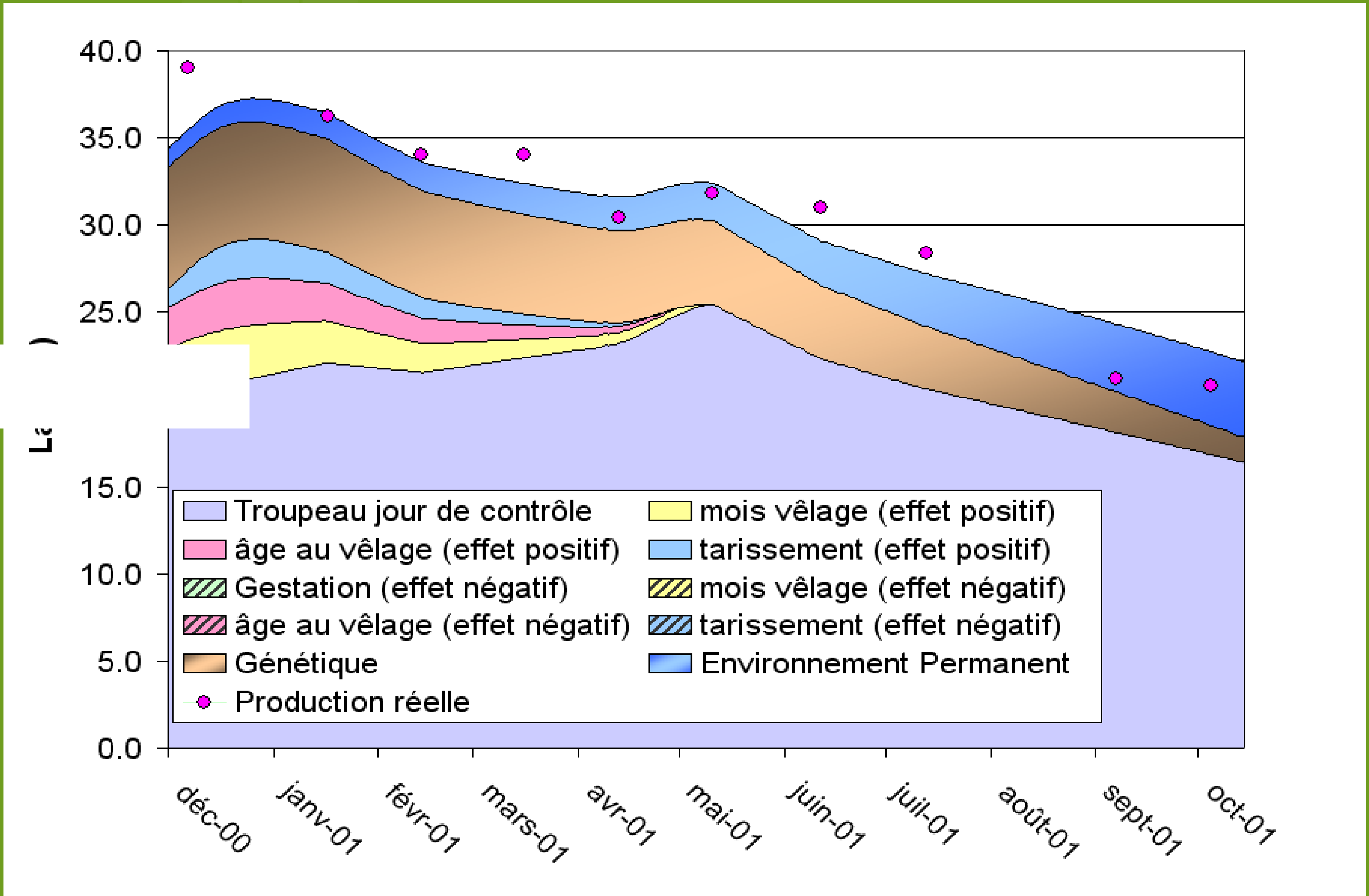


Large difference between Normande *vs* Montbéliarde and Holstein

Prediction of expected production of a cow



Comparision between predicted and observed values



test day model implementation: on milk production data from BAIF field

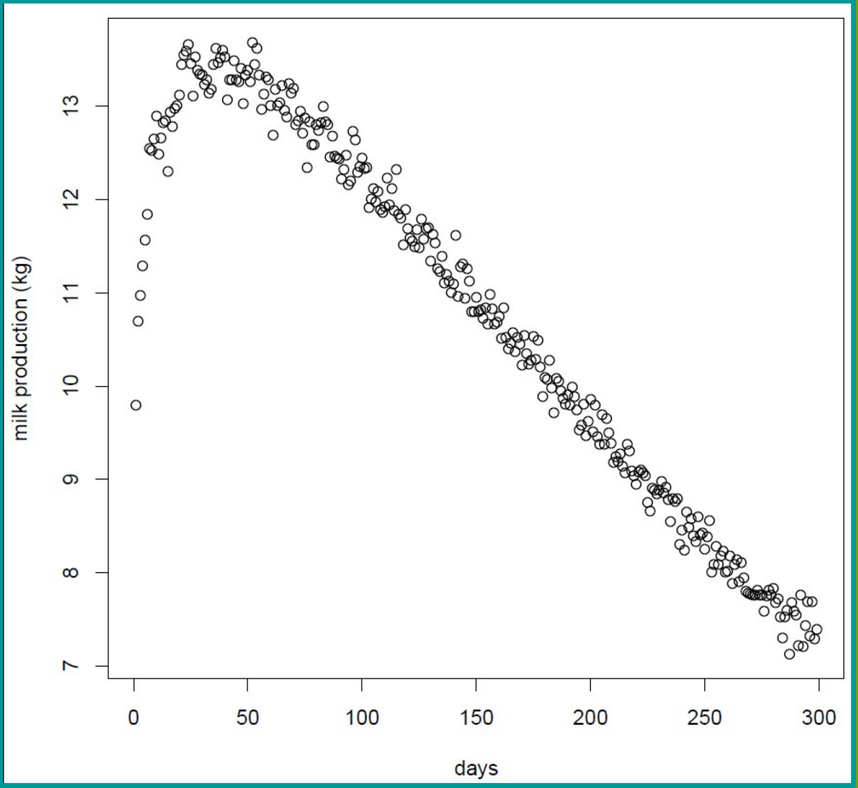
Clotilde Patry (2012)

Data available

- BAIF field data: collected during milk recording (did not belong to progeny test)

- Performances for 6,675 crossbred cows
  - Holstein or Jersey crossbreds
  - Recorded from 1994 to 2010
  - >120,000 observations
  - Up to the 6th lactation rank (mostly 1st or 2nd lactation)

Average milk production per day



Mean = 10.5 kg / day
Range: from 1 to 34 kg / day

Effects included in the model

FIXED

- Contemporay group : « Center – Month – Year » instead of « herd – test date »
- Genetic type (% crossbreeding)
- **1 Lactation curve per lactation number**

RANDOM

- Permanent environment effect (cow x lactation number) : repeatability

RANDOM

- Genetic effect of interest: **sire effect**

- Main problem: many recorded cows with unknown sire

Data analysis: a 2-step model

- **STEP 1: consider all HF crossbreds**

➔ Estimation of the contemporary group effect

+ permanent environment effect

- **STEP 2: Only records of daughters from sires with > 5 daughters**

➔ Genetic evaluation based on records corrected for the contemporary group and permanent environmental effects

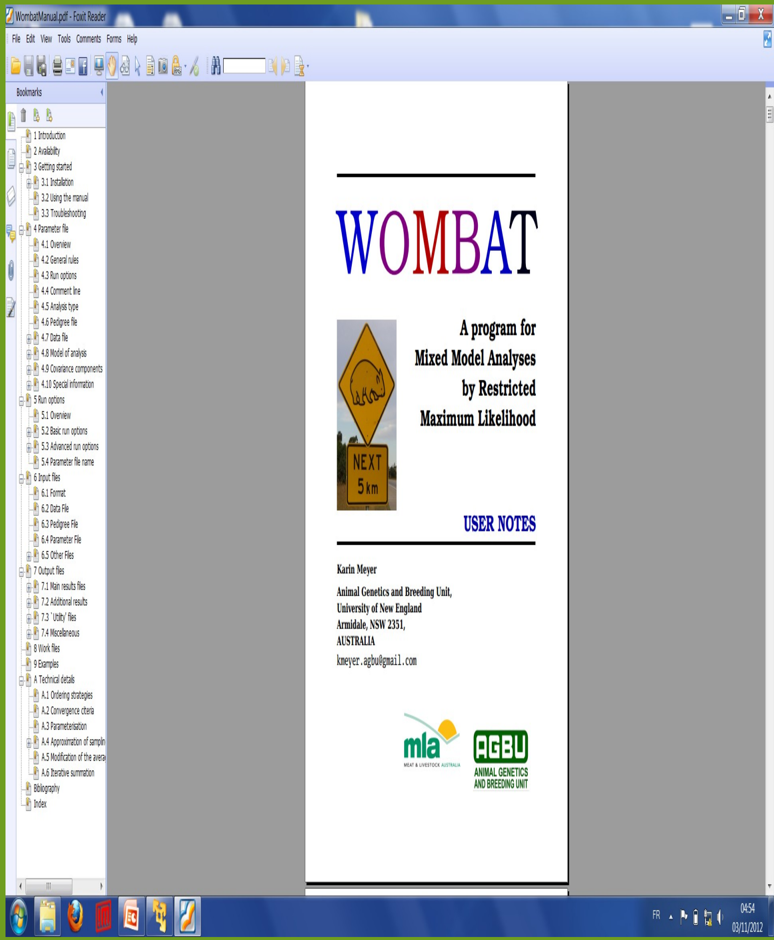
 Get estimated breeding values for sires

Steps to implement for routine evaluations

Integrated and automatic procedure

Programing language and software

- Tool based on free software :
  - **R** for the statistics, curves
  - **WOMBAT** (Karin Meyer) for the genetic parameter estimation and breeding values estimations
- Tool based on free language for programming:
  - **AWK:** to edit the data
  - **Shell script** : to go from one step to another => integrated steps
  - ⇒ Well documented software and languages, available on the web
  - ⇒ Documented application: **to make it easier to transfer**



www/didgeridoo.une.edu.au/km/wombat.php

Illustration from BAIF field data

- **Data editing**

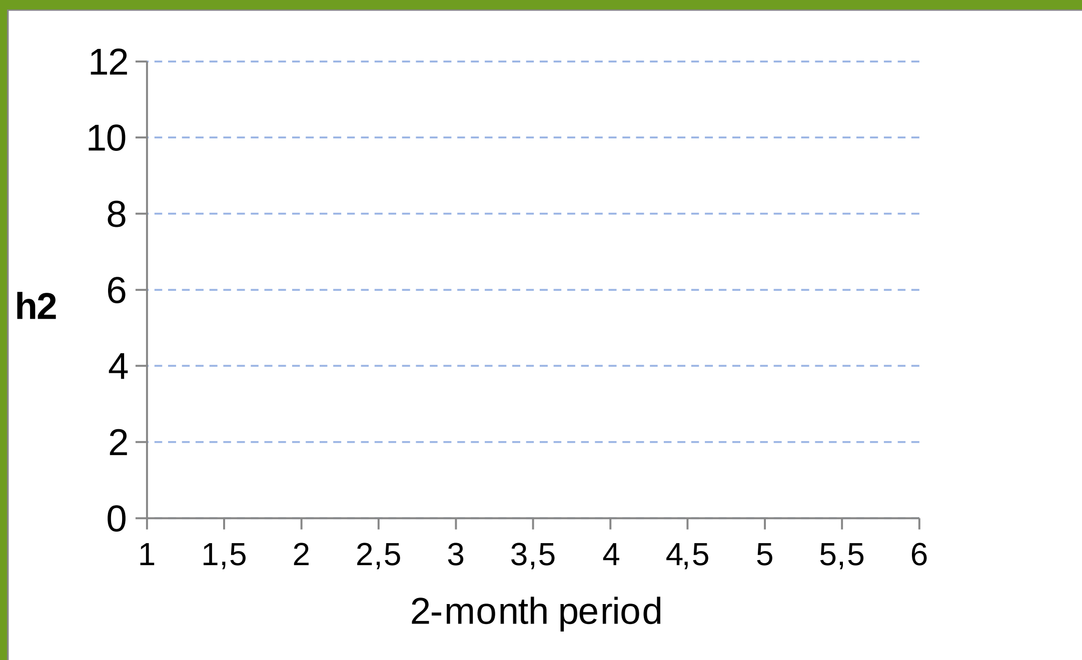
Editing steps	*# test day records*	*# cows*	*# sires*
After deletion for inconsistencies, missing information	*~125,000*	*6,675*	*> 400*
> 5 records / lactation	*~120,000*	*~6,600*	*> 400*
> 5 daughters/sire	*~38,000*	*>2,200*	*95*

-4%

-70%

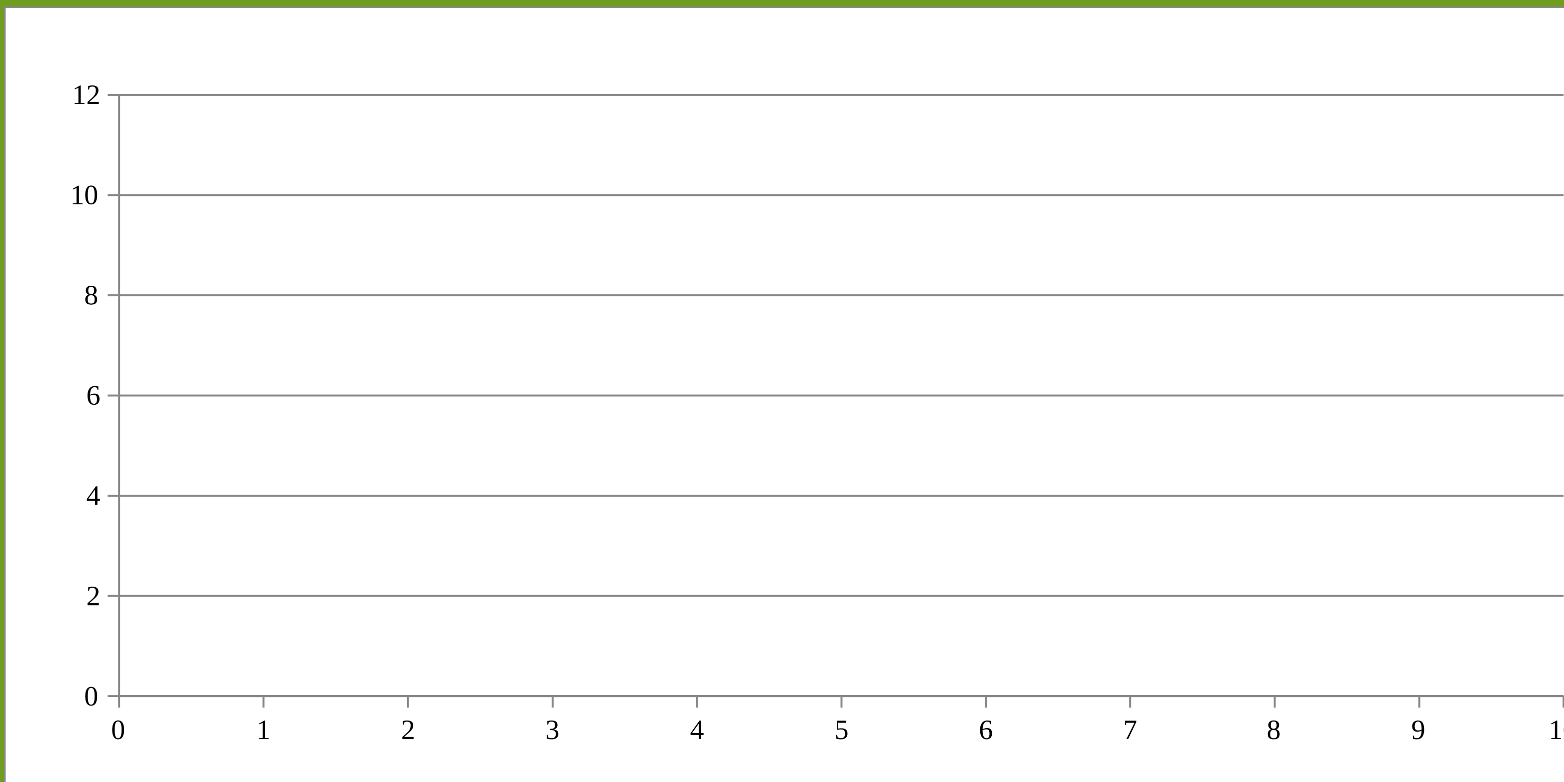
Illustration from BAIF field data

- Estimation of heritability through lactation using 3 models for lactation curve
  - Fixed regression (constant genetic effect) (wo)
  - Legendre polynomials – 3rd order (leg 3)
  - Splines – 3rd order (could have been done with Wilmink too)



Variances

- Genetic parameter estimation:



Month in milk

Genetic evaluation

- Estimation of breeding values (test day model, splines)
- 192 candidates
- Distribution:
  - Min(EBV)= - 4.49
  - Max(EBV)= 4.5
  - Mean(EBV)=0