KU LEUVEN

FACULTEIT BIO-INGENIEURSWETENSCHAPPEN

Genetic parameters for scrotal circumference and weight traits in Black Hereford cattle

Genetische parameters voor scrotumomtrek en gewichtseigenschappen bij Black Hereford runderen

Promotoren:
Dr. ir. Steven Janssens
Prof. Nadine Buys
Afdeling Dier en Mens
Departement Biosystemen

Masterproef voorgedragen
tot het behalen van het diploma van
Master of science in de bio-ingenieurswetenschappen:
landbouwkunde

Roel Meyermans

juni 2017

KU LEUVEN

FACULTEIT BIO-INGENIEURSWETENSCHAPPEN

Genetic parameters for scrotal circumference and weight traits in Black Hereford cattle

Genetische parameters voor scrotumomtrek en gewichtseigenschappen bij Black Hereford runderen

Promotoren:
Dr. ir. Steven Janssens
Prof. Nadine Buys
Afdeling Dier en Mens
Departement Biosystemen

Masterproef voorgedragen
tot het behalen van het diploma van
Master of science in de bio-ingenieurswetenschappen:
landbouwkunde

Roel Meyermans

juni 2017

Acknowledgements

I would like to thank my promotor and supervisor Dr. ir. Steven Janssens. The door to his office was always open. His advice was always useful and he steered me in the right direction during my research and writing. I also would like to thank my co-promotor Prof. Dr. Nadine Buys for giving me the chance to work out my thesis at the Livestock Genetics research group and for offering the place to work at the lab. Thanks to Katrien W., Liesbeth, Susanne, Wim and Katrien V. for the nice working atmosphere at the office.

I am grateful to Dr. Ignace Moyaert and Prof. Dr. Peter Goos for their participation in my jury.

My sincere thanks to Joe and Norma Hoagland for inviting me at their ranch in Kansas. Your hospitality is truly amazing, you really made me feel at home. It was a great experience being there! Thank you for providing the data for my research and introducing me to the Black Hereford breed.

I would also like to thank my friends for the five amazing years we had here in Leuven!

Last but not least, I would like to thank Zoë, my parents and my family for supporting me during my years of study and through the process of this thesis.

Abstract

The scrotal circumference of yearling bulls and its relation to productive and reproductive traits has been previously studied in major commercial beef breeds. This study evaluates the heritability of yearling scrotal circumference and its relation to weight traits in Black Hereford cattle.

Records for yearling scrotal circumference (YSC; n = 894), birth weight (BW; n = 2350), weaning weight (WW; n = 2077) and yearling weight (YW; n = 1434) were obtained in Black Hereford cattle. Genetic parameters were estimated using an animal model. The final model for YSC included the age at measurement as covariate and fixed effects of birth year, the observer and the age of dam. The final models for BW, WW and YW included the age at measurement (if available) as covariate and fixed effects of birth year, age of dam, sex and embryo transfer. Random effects in the four models were additive genetic effects. The four-trait animal model was fitted by best linear unbiased prediction with restricted maximum likelihood (co)variance estimates. A univariate model for yearling scrotal circumference was also estimated by a Markov chain Monte Carlo algorithm.

Heritabilities were estimated at 0.49, 0.43, 0.47 and 0.45 for YSC, BW, WW and YW, respectively. These estimates are moderate to high and indicate that selection for these traits is possible in Black Herefords. Genetic correlations of 0.22, 0.54 and 0.59 were found between YSC and BW, WW and YW, respectively. The genetic correlation between WW and YW was estimated at 0.89. The relation of YSC to the age at first calving in the bulls daughters could not be studied due to less consistent data of the age at first calving. An increase of YSC and weight traits was found in bulls of older dams. Expected progeny differences for YSC, BW, WW and YW were proposed for 4099 Black Herefords. These expected progeny differences can be included in a breeding program.

Samenvatting

De scrotumomtrek bij jaarling stieren en de relatie hiervan met productie- en reproductiekenmerken is reeds vaak bestudeerd in verschillende commercieel belangrijke vleesveerassen. Deze studie evalueert de erfelijkheidsgraad van de jaarling scrotumomtrek en de relatie tussen de jaarling scrotumomtrek en gewichtskenmerken in het Black Hereford runderras.

Jaarling scrotumomtrek (n = 894), geboortegewicht (n = 2350), speengewicht (n = 2077) en jaarlinggewicht (n = 1434) observaties werden verkregen voor 2532 Black Hereford runderen. Er werden genetische parameters geschat voor deze vier kenmerken door middel van een diermodel. In dit diermodel werden geboortejaar, leeftijd van de moeder, geslacht, waarnemer en embryo transfer als vaste effecten opgenomen, de leeftijd op moment van meting als covariabele en een additief genetisch effect als random effect. Het diermodel werd gefit door middel van best linear unbiased prediction met restricted maximum likelihood (co)variantie schattingen. Daarnaast werd het univariaat model voor jaarling scrotumomtrek ook geschat door middel van een Markov chain Monte Carlo algoritme.

De erfelijkheidsgraden voor jaarling scrotumomtrek, geboorte-, speen- en jaarling gewicht werden geschat op respectievelijk 0.49, 0.43, 0.47 en 0.45. Er werden genetische correlaties tussen jaarling scrotumomtrek en geboorte-, speen- en jaarling gewicht geschat van respectievelijk 0.22, 0.54 en 0.59. De genetische correlatie tussen speen- en jaarling gewicht werd geschat op 0.89. Er werd een toename vastgesteld voor jaarling scrotumomtrek naarmate de moederdieren ouder zijn. De relatie tussen jaarling scrotumomtrek en de leeftijd van eerste kalving in vaarzen kon niet worden onderzocht door moeilijkheden bij de berekening van deze leeftijd. Bovendien werden er fokwaarden voor jaarling scrotumomtrek, geboorte-, speen- en jaarling gewicht berekend voor 4099 Black Hereford runderen. Deze fokwaarden kunnen gebruikt worden in het fokprogramma voor het ras.

List of abbreviations

ABHA	American Black Hereford Association
AFC	Age at first calving
AIC	Akaike information criterion
AP	Age at puberty
BLUE	Best linear unbiased estimate
BLUP	Best linear unbiased prediction
BW	Birth weight
FSH	Follicle stimulating hormone
EBV	Estimated breeding value
EPD	Expected progeny difference
ET	Embryo transfer
LH	Luteinizing hormone
MB	Mean bias
MCMC	Markov chain Monte Carlo
MME	Mixed model equations
MSB	Mean squared bias
PEV	Prediction error variance
WA	Weaning age
WW	Weaning weight
YA	Yearling age
(Y)SC	(Yearling) scrotal circumference
YW	Yearling weight

List of used symbols

The following list is the list of used symbols in the literature, material and methods and results sections, in order of appearance.

p	The measured phenotype
g	The genetic aptitude
E	The environmental effect
a	The additive genetic effect (equal to the true breeding value of the animal)
d	The dominance effect
i	The genetic interaction effect
σ_p^2	The phenotypic variance
σ_a^2	The additive genetic variance
σ_d^2	The variance due to dominance effects
σ_i^2	The variance due to interaction effects
σ_E^2	The variance due to environmental effects
σ_g^2	The (total) genetic variance
h^2	The (additive) heritability of a trait
n	The number of observations
$\boldsymbol{Y}$	The vector of dependent variables
$\boldsymbol{\beta}$	The fixed effects vector
$\boldsymbol{a}$	The additive genetic effects vector
$\boldsymbol{e}$	The residual effects vector
$\boldsymbol{X}$	The fixed effects design matrix
$\boldsymbol{Z}$	The random effects design matrix
σ_e^2	The residual variance
$\boldsymbol{R}$	The residual covariance matrix
$\boldsymbol{G}$	The genetic covariance matrix
$\boldsymbol{A}$	The relationship matrix
$\widehat{\boldsymbol{a}}$	The BLUP of **a** (the predicted breeding value)
$\widehat{\boldsymbol{\beta}}$	The BLUE of $\boldsymbol{\beta}$
$\boldsymbol{V}$	The (total) covariance matrix
α	The ratio of residual variance over the additive genetic variance
$\boldsymbol{W}$	The common environmental design matrix

$\boldsymbol{c}$	The common environmental effects vector
r_i	The accuracy value of the EBV/EPD for animal *i*
d_i	The ith diagonal element of the inverse of $\boldsymbol{Z'Z} + \boldsymbol{A^{-1}}\alpha$
r_s	Accuracy value for the EBV/EPD of the sire of an animal
r_d	Accuracy value for the EBV/EPD of the dam of an animal
$L(\boldsymbol{Y})$	Fishers likelihood function of $\boldsymbol{Y}$
$p(\ )$	A probability
V	The variance limit of the inverse-Wishart distribution
nu	The degree of belief parameter of the inverse-Wishart distribution
CR	The correlated response
$\Delta G_{2(1)}$	The genetic progress of trait 2 by selection for trait 1
h_1	The reliability of trait 1 (=$\sqrt{h_1^2}$)
h_2	The reliability of trait 2 (=$\sqrt{h_2^2}$)
I	The selection intensity
$r_{g_{12}}$	The genetic correlation between trait 1 and 2
AF	Adjustment factor for age
AF	Adjustment factor for age of dam
R^2	The coefficient of determination
y_i	The dependent variable for observation *i*
$\hat{y}_i$	The expected value for the dependent variable of observation *i*
$r_{\ g\ (mn)}$	The genetic correlation coefficient between traits *m* and *n*
$\sigma^2_{a\ (mn)}$	The additive genetic covariance between traits *m* and *n*
$\sigma^2_{a\ (n)}$	The additive genetic variance for trait *m*
$\sigma^2_{a\ (m)}$	The additive genetic variance for trait *n*
k	The number of estimable parameters of the model

List of tables

Table 2.1 Two hypothetical bulls and their expected progeny differences for birth weight, weaning weight and calving ease 5

Table 5.1 Age adjustment factors for YSC in yearling bulls between 300 and 400 days of age by different authors for Hereford and Angus cattle 27

Table 5.2 Scrotal circumference adjustment factors for age of dam 27

Table 5.3 Overview of yearling scrotal circumference measurements in different studies on different beef breeds. The mean YSC and the standard deviation of YSC are given together with the heritability estimate for YSC or age-adjusted YSC 28

Table 5.4 Several heritability estimates for age at puberty and genetic correlations between a sires yearling scrotal circumference and age at puberty in daughters in different beef breeds 30

Table 5.5 Several heritability estimates for age at first calving and genetic correlations between a sires yearling scrotal circumference and age at first calving in daughters in different beef breeds 31

Table 6.1 A list of heritability estimates for birth weight, weaning weight and yearling weight in Angus, Hereford and Black Hereford cattle herds 35

Table 6.2 A list of genetic correlations found in literature between for birth weight, weaning weight and yearling weight in Angus and Hereford cattle 35

Table 7.1 List of variables in the performance file of the Black Hereford cattle 36

Table 7.2 List of variables in the pedigree file of the Black Hereford cattle 36

Table 8.1 List of the variables available for analysis 41

Table 8.2 List of the number of observations for every level of the variables for yearling scrotal circumference, birth weight, weaning weight and yearling weight 42

Table 8.3 The descriptive statistics for yearling scrotal circumference, birth weight, weaning weight, yearling weight, weaning age and yearling age measured in Black Hereford cattle from 1997 to 2016 43

Table 8.4 The Pearson correlations between yearling scrotal circumference, birth weight, weaning weight and yearling weight, weaning age and yearling age measured in Black Hereford cattle from 1997 to 2016 43

Table 8.5 Summary of the significant effects in univariate least squares regression models with yearling scrotal circumference, birth weight, weaning weight and yearling weight as dependent variable 46

Table 8.6 Overview of the composition of the five animal models used in this analysis 48

Table 8.7 AIC values and the R^2 values for yearling scrotal circumference, birth weight, weaning weight and yearling weight for the five models 49

Table 8.8 Mean bias and mean squared bias of the five models for yearling scrotal circumference, birth weight, weaning weight and yearling weight 49

Table 8.9 The Spearman correlations of the expected progeny differences between the five models for yearling scrotal circumference, birth weight, weaning weight and yearling weight. The upper diagonal matrix gives the correlations for the whole population and the lower diagonal matrix those for the active sire population 50

Table 8.10 Pearson correlations between the observed and the estimated observations and the estimated observations mutually between yearling scrotal circumference, birth weight, weaning weight and yearling weight for the five models. 51

Table 8.11 Variance and covariance estimates for yearling scrotal circumference, birth weight, weaning weight and yearling weight estimated by model 3 52

Table 8.12 Genetic correlations, heritabilities and the residual variances for yearling scrotal circumference, birth weight, weaning weight and yearling weight estimated by model 3 52

Table 8.13 An excerpt of the EPD list, giving the EPDs for yearling scrotal circumference, birth weight, weaning weight and yearling weight, the respective accuracy values and the number of offspring for eight Black Hereford animals, based on the estimates of model 3 55

Table 8.14 Overview of the mean, variance, standard deviation and range of the EPDs for yearling scrotal circumference, birth weight, weaning weight and yearling weight, estimated by model 3 55

Table 8.15 The Spearman correlations between the EPDs for yearling scrotal circumference, birth weight, weaning weight and yearling weight, as estimated by model 3 55

Table A5.1 Overview of the linear regression results FOR YEARLING SCROTAL CIRCUMFERENCE, BIRTH WEIGHT, WEANING WEIGHT AND YEARLING WEIGHT xi

Table A5.2 Summary of the number of observations, adjusted R² and root mean squared error for the univariate linear regression models of yearling scrotal circumference, birth weight, weaning weight and yearling weight xii
Table A6.1 Variance and covariance estimates for yearling scrotal circumference, birth weight, weaning weight and yearling weight estimated by the five models. xiii
Table A6.2 The estimated genetic correlations, residual correlations and heritabilities in the five models xv
Table A7.1 Overview of the parameter estimates of the five models for yearling scrotal circumference xvi
Table A7.2 Overview of the parameter estimates of the five models for birth weight xvii
Table A7.3 Overview of the parameter estimates of the five models for weaning weight xviii
Table A7.4 Overview of the parameter estimates of the five models for yearling weight xix
Table A8.1 Overview of the posterior means of the approximated posterior density functions, the 95% posterior density intervals and the effective sample sizes for the effects included in the animal model with yearling scrotal circumference as dependent variable xx

List of figures

Figure 2.1 An example of the use of expected progeny differences 5

Figure 2.2 An example trace from Hadfield (2016) where the burn-in period is indicated. After the burn-in period, the Markov chain converges to a value of ± 7.5. 13

Figure 2.3 The probability density function of an inverse-gamma function with V=1 and nu varying from 0.002 to 1 14

Figure 3.1 An Angus bull 17

Figure 3.2 A Hereford bull 18

Figure 3.3 A Black Hereford bull 20

Figure 3.4 The crossing scheme that has been used to create the Black Hereford breed 21

Figure 5.1 Left: A scrotal tape, used for the scrotal circumference measurement. Right: The use of a scrotal tape in practice on a yearling Black Hereford bull 26

Figure 8.1 Scatterplot matrix for yearling scrotal circumference, birth weight, weaning weight, yearling weight, weaning age and yearling age, measured in Black Hereford cattle from 1997 to 2016 44

Figure 8.2 Frequency histogram of the age at first calving of cows in the Black Hereford registry from 1992 to 2016 45

Figure 8.3 Plot of the age of dam at first calving by birth year in the Black Hereford registry from 1992 to 2016 45

Figure 8.4 The additive effect of the age of dam on v as estimated by model 3 54

Figure 8.5 The additive effect of the year of birth on yearling scrotal circumference, birth weight, weaning weight and yearling weight as estimated by model 3 54

Figure 8.6 Frequency historgrams of the yearling scrotal circumference, birth weight, weaning weight and yearling weight EPDs estimated by BLUP 56

Figure 8.7 The mean of the posterior distribution of the age of dam on yearling scrotal circumference, relative to 2-year old dams, as estimated by the MCMC algorithm 58

Figure 8.8 Frequency historgram of the yearling scrotal circumference, estimated by MCMC 58

Table of content

Acknowledgements i
Abstract ii
Samenvatting iii
List of abbreviations iv
List of used symbols v
List of tables vii
List of figures ix
Table of content x
1 Introduction and research objective 1
2 Quantitative genetics 2
  2.1 Introduction to quantitative genetics 2
  2.2 The heritability of a trait 3
  2.3 Breeding values and expected progeny differences 4
    2.3.1 Estimated breeding values 4
    2.3.2 Expected progeny differences 4
    2.3.3 The animal model and best linear unbiased prediction 6
    2.3.4 The accuracy value of an EPD 9
    2.3.5 Restricted maximum likelihood 10
    2.3.6 The Markov chain Monte Carlo method 11
    2.3.7 The differences between MCMC and REML 14
  2.4 The phenotypic and genetic correlation of traits 15
  2.5 Crossbreeding 16
3 The Black Hereford breed and its origin 17
  3.1 The Angus breed 17
  3.2 The Hereford breed 18
  3.3 Commercial cattle crossbreeding 19
  3.4 The Black Hereford breed 20
  3.5 The black coat color 22
4 The reproduction of cattle 23
  4.1 The puberty of cattle 23
  4.2 The gestation and the optimal age of calving 24
5 Scrotal circumference 26
  5.1 Scrotal circumference 26
  5.2 Yearling scrotal circumference measurements and heritability estimates 27
  5.3 Scrotal circumference and production traits 29

5.4 Scrotal circumference and male reproduction traits 29
5.5 Scrotal circumference and female reproduction traits 30
5.5.1 The age at puberty 30
5.5.2 The age at first calving 31
5.6 Scrotal circumference EPDs and selection 32
6 Weight traits in beef cattle 34
6.1 Birth weight, weaning weight and yearling weight 34
6.2 The heritability of weight traits 35
6.3 The genetic correlations between weight traits 35
7 Material and methods 36
7.1 Material 36
7.1.1 Observation records 36
7.1.2 Pedigree records 36
7.2 Methods 37
7.2.1 Data management and statistical analysis 37
7.2.2 BLUP and REML estimation 37
7.2.3 The MCMC algorithm 37
7.2.4 Computer usage 38
7.2.5 Formulas used in validation and computation 38
8 Results 41
8.1 Data management and descriptive statistics 41
8.2 The age at first calving 45
8.3 Dams and sires versus birth year 46
8.4 Linear regression results 46
8.5 The animal model results from airemlf90 47
8.5.1 Description of the models 47
8.5.2 Model validation 49
8.5.3 Estimates of variances, covariances, genetic correlations and heritabilities 52
8.5.4 Parameter estimates 53
8.5.5 EPDs 55
8.6 The MCMC algorithm 57
9 Discussion 59
9.1 Data management and statistics 59
9.2 Age at first calving 59
9.3 Dam and sires vs birth year 60
9.4 Linear regression results 60
9.5 Animal model results 61
9.5.1 Model description 61

9.5.2 Validation and selection of the different models 61
9.5.3 Estimates of variances, covariances, genetic correlations and heritabilities 63
9.5.4 Parameter estimates 64
9.5.5 EPDs 66
9.5.6 A remark on model 1 67
9.6 The MCMC algorithm 68
10 Conclusion 70
11 References 72
Vulgarizing summary 77
Appendix 1 – The J&N Ranch management i
Appendix 2A – BLUP and the MME ii
Appendix 2B – BLUP and accuracy values v
Appendix 3 – Bayesian statistics: Gibbs sampling vi
Appendix 4 – Parameter file for renumf90 x
Appendix 5 – Results of the linear regression of the four traits xi
Appendix 6 – (Co)variance components obtained using REML xiii
Appendix 7 – Parameter estimates xvi
Appendix 8 – Output of the MCMC algorithm xx

1 Introduction and research objective

The Black Hereford is a recent beef breed and increases in popularity. It arose from crossings between Angus and Hereford cattle and combines the black coat from Angus with the feed efficiency of Hereford. After focusing on meat quality and production traits, breeders want to improve the reproductive traits of the Black Hereford.

One of these points of improvement is the age at first calving of heifers. In optimal commercial conditions, the age at first calving is 24 months. In other beef breeds, this age at first calving is found to be a lowly heritable trait and thus direct selection for this trait would be slow. Nevertheless, the age at first calving is in other breeds found to be genetically correlated to the yearling scrotal circumference of the heifers father.

The general objective of this study is to examine the relation between yearling scrotal circumference and weight- and reproductive traits. More in detail, the study will include:

- The estimation of the heritability for yearling scrotal circumference in the Black Hereford breed, based on the available data.
- The estimation of the heritabilities for birth -, weaning - and yearling weight.
- The calculation of genetic correlations between yearling scrotal circumference, age at first calving, birth -, weaning - and yearling weight.
- A proposal for breeding values for yearling scrotal circumference. These expected progeny differences can be used in a breeding program in order to decrease the age at calving in heifers.
- A proposal for breeding values for birth -, weaning - and yearling weight.

The literature review begins with an introduction to quantitative genetics and animal models. Next, the Black Hereford breed is discussed, followed by a section about the reproduction of cattle. Then, the scrotal circumference trait and its relevance to other production and reproduction traits is discussed, followed by a description of weight traits in beef cattle.

Literature

2 Quantitative genetics

2.1 Introduction to quantitative genetics

The genotype of an animal is defined as the total of hereditary information present in the genome of the animal (Nicholas, 2010). This information is present in the form of genes. The phenotype is defined as the total of the actual observed characteristics and traits of the animal. Depending on the trait, the phenotype is caused by a single or by multiple genes. The traits influenced by only one gene are called monogenic. An example of a monogenic trait in cattle is the presence of horns (Nicholas, 2010). Traits which are influenced by a lot of separate genes are called polygenic or quantitative traits. Most of the observed phenotypes in cattle, such as growth, fertility, longevity and height, are polygenic. For these traits, dozens or even hundreds of genes work together to form the phenotype. The effects of these genes are often additive and cause a continuous variation in the observed phenotype (Lynch & Walsh, 1998; Nicholas, 2010).

For almost every polygenic trait, the observed phenotype is influenced by environmental factors. These environmental factor are typically feed, weather conditions and different management decisions. This results in the fact that the observed phenotype of an animal does not directly reflect its genetic value. The genetic value of an animal has to be estimated from the measured phenotypic values. Therefore, observed phenotypes can be described by the following additive model (Acquaah, 2012; Nicholas, 2010)

$$p = g + E + g * E \tag{2.1}$$

where p is the measured phenotype, g the genetic aptitude, E the total of environmental factors and $g * E$ the interaction between genetic and environmental factors. In most models, the term $g * E$ is neglected to make computations more feasible. It should be noted that in some cases this term is not neglectable (e.g. when animals are genetically adapted to a certain environment). The genetic aptitude can be subdivided into (Nicholas, 2010):

$$g = a + d + i \tag{2.2}$$

with a the additive genetic component, d the component due to dominance effects and i the factor due to interactions between multiple genes, called epistatic effects (Lynch & Walsh, 1998; Nicholas, 2010).

The phenotypic variance (σ_p^2) is the extent to which animals differ in their phenotypic values. Given (2.1 and 2.2), this phenotypic variance can be subdivided in (Acquaah, 2012; Nicholas, 2010):

$$\sigma_p^2 = \sigma_a^2 + \sigma_d^2 + \sigma_i^2 + \sigma_E^2 \tag{2.3}$$

where σ_a^2 is the additive genetic variance, which is equal to the variance in breeding values, σ_d^2 is the variance due to dominance effects, σ_i^2 is the variance in effects due to epistatic interactions and σ_E^2 is the variance in non-genetic (environmental) effects. The term $\sigma_d^2 + \sigma_i^2$ is called the non-additive genetic variance. The sum of the additive- and non-additive genetic variances is called the (total) genetic variance (σ_g^2) (Acquaah, 2012; Nicholas, 2010)(Nicholas, 2010).

2.2 The heritability of a trait

For a polygenic trait, the ratio of the additive genetic variance to the phenotypic variance is called the (additive) heritability (h^2) of the trait. It can be calculated as (Acquaah, 2012; Nicholas, 2010):

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_d^2 + \sigma_i^2 + \sigma_E^2} \tag{2.4}$$

Since the heritability is a ratio, it always varies between zero and one. Notice that a heritability of zero does not necessarily mean the trait is not genetically determined, it only indicates that there is no observed genetic variance.

The heritability of a trait is population dependent and can change over time. It can also change by selection: when selection decreases the observed genetic additive variance, the heritability will decrease. Also changes and evolutions in management conditions, feeding or measuring techniques can influence the estimation of the heritability of a trait (Acquaah, 2012; Nicholas, 2010).

Some traits can have a high heritability (typically > 0.55), which means a large part of the observed variance can be attributed to the genetic variance. These traits are typically performance- and production traits. A high heritability indicates that the offspring tends to perform similar to their parents. Traits with a lower heritability (typically < 0.20) indicate a high influence of environmental variance in comparison to the genetic variance. Traits with low heritability are typically survival- and reproductive traits, like calving ease in cattle (Bullock, 2009; Nicholas, 2010; Simm, 1998).

2.3 Breeding values and expected progeny differences

2.3.1 Estimated breeding values

The fact that for many traits the genotype cannot be determined directly makes it nearly impossible to identify an animal as a good parent 'by eye'. In order to know what the animal will pass to its progeny by genetics, it is necessary to determine the true breeding value of the animal. Since it is not possible to see the true breeding value based on the phenotype, it needs to be estimated from phenotypic traits. This estimate is called the estimated breeding value (EBV) (Nicholas, 2010).

2.3.2 Expected progeny differences

Another form of expressing EBVs for an animal, is by using expected progeny differences (EPDs). This EPD shows the estimated genetic value that the animal will pass to its offspring. As each parent contributes 50 % of its genes to its progeny, an EPD is half the value of the EBV. The EBV of the progeny is the sum of the EPDs of both parents. It is important to note that EPDs predict the expected difference in performance in comparison to a reference population, not the actual performance of an animal (Greiner, 2009).

In the USA, most beef breed associations will advertise their cows and bulls by using EPDs instead of using EBVs. Most cattle herdbooks report EPDs for birth weight, weaning weight, yearling weight and calving ease. Some larger herdbooks estimate EPDs for 22 traits (American Angus Association, 2016c).

One essential component of a consistent breeding program is the breeding goal. This breeding goal defines what the breeders want to achieve with their herd or breed. The breeders will decide which traits will be important for their cattle and which traits will arm their breed for future challenges. EBVs or EPDs for these traits will help the breeders to make optimal choices in their breeding program.

EPDs (and EBVs) are always compared to a given genetic base reference. The genetic base can be chosen freely (e.g. a historic level). Therefore, EPDs are always expressed as the difference (+ or -) between an individual animal and the chosen baseline. For different breeds, the genetic bases might differ. Because of that, it is not useful to compare EPDs of animals from a different breed. To do so, there is a need for breed EPD adjustment factors (Nicholas, 2010).

The differences in EPDs between two animals predict differences in performance between their future offspring when each of these two animals is mated to animals of similar average genetic merit. The use of EPDs is illustrated in the following example. The EPDs for two different hypothetical bulls are given in the following table (2.1).

Table 2.1 Two hypothetical bulls and their expected progeny differences (EPD) for birth weight (BW), weaning weight (WW) and calving ease (CE)

	BW EPD (kg)	WW EPD (kg)	CE EPD (% of unassisted births)
Bull A	+2.3	+8	+1
Bull B	+1	+4	+4

Assume that each bull will be mated to a similar set of cows. The difference in birth weight EPDs between bull A and bull B is 1.3 kg. Thus, it is expected that calves from bull A would be on average 1.3 kg heavier at birth than calves sired by bull B. The same applies to the weaning weight EPDs: calves from bull A are expected to weigh 4 kg more at weaning than calves sired by bull B. For calving ease, it is expected that calves from bull B would experience 3% more unassisted births than bull A. So it is expected that heifers have less calving problems when mated to bull B. The given example is illustrated in the following figure (2.1), where both bulls are mated to a cow with +1 kg, +3 kg and +2 % as birth weight, weaning weight and calving ease EPDs, respectively.

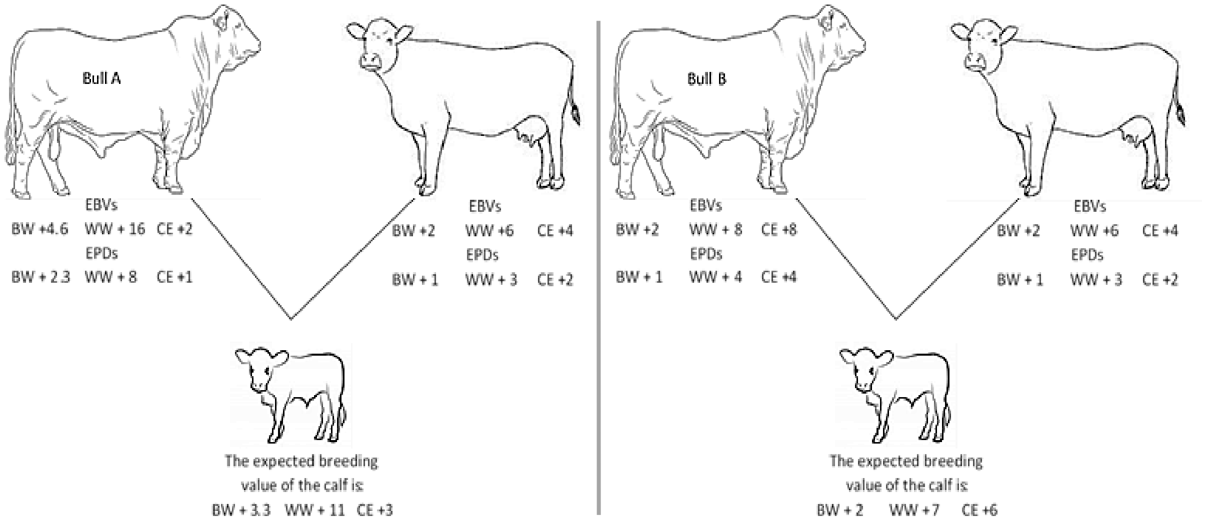

Figure 2.1 An example of the use of expected progeny differences. On the left side the hypothetical bull A, on the right bull B with EPDs for birth weight (BW) (in kg), weaning weight (WW) (in kg) and calving ease (CE) (in % of unassisted births)

2.3.3 The animal model and best linear unbiased prediction

EBVs, and thus EPDs as well, are currently often calculated by an animal model. This model uses the information of the measurements on the animal itself, its relatives and other herd mates. This animal model requires pedigree information and phenotypic data of the animals. The pedigree information gives the genetic relationship of each animal in the dataset to the other animals in the dataset. This way, performances from ancestors and descendants can be used to calculate the EBVs (Nicholas, 2010).

The animal model is a mixed model, composed of random and fixed effects (Mrode, 1996). Random effects usually follow a Normal distribution around the mean value with a certain variance. Typical random effects in the animal breeding context are breeding values. For a fixed effect, the levels will be estimated as deviation from the mean value. These estimates are called the best linear unbiased estimates (BLUE). Common fixed effects are sex, nutrition levels and the age of the mother (also called dam). Each EBV calculation involves solving a set of equations. The more data and pedigree information available, the better the EBV will approach the true breeding value of the animal. The animal model can correct for assortative mating (this happens when the breeder uses his best bull to sire his best cows), selection and inbreeding, if this information is included in the model (Bullock, 2009; Beef Improvement Federation, 2010; Nicholas, 2010).

The animal model relies on work of Henderson (1949) who developed a methodology called best linear unbiased prediction (BLUP) (Mrode, 1996). His theory on mixed models found usage in genetic evaluation of livestock animals because of its desirable statistical properties. At first, BLUP was used in sire models but the availability of more computing power has made it possible to estimate more advanced models, like the animal model. By using BLUP, it is possible to predict breeding values and estimate fixed effects simultaneously. The name 'BLUP' points out some key features of the method (Mrode, 1996):

- Best: the correlation between the true ($\boldsymbol{a}$) and predicted breeding value ($\widehat{\boldsymbol{a}}$) is maximized. In other words, the prediction error variance $\left(PEV = var(\boldsymbol{a} - \widehat{\boldsymbol{a}})\right)$ is minimized. This PEV can be interpreted as the fraction of additive genetic variance that is not accounted for by the prediction.
- Linear: the predictors are linear functions of observations.
- Unbiased: the estimations of random variables, such as animal breeding values, and of estimable functions of fixed effects are unbiased ($E(\boldsymbol{a}|\widehat{\boldsymbol{a}}) = \widehat{\boldsymbol{a}}$)**.**
- Prediction: the method involves prediction of true breeding values.

Mixed linear models have the following general (matrix) form (Schaeffer, 2000; Mrode, 2005):

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{e} \tag{2.5}$$

where:

- $\boldsymbol{Y}$ is a n x 1 vector of observations with n the number of observations of the measured trait(s)
- $\boldsymbol{\beta}$ is a p x 1 vector of fixed effects with p the number of levels of the fixed effects
- $\boldsymbol{a}$ is a q x 1 vector of random additive genetic effects with q the number of levels of random effects
- $\boldsymbol{e}$ is a n x 1 vector of random residual effects, attributed to every observation
- $\boldsymbol{X}$ is a design matrix of order n x p, which relates observations to their fixed effects
- $\boldsymbol{Z}$ is a design matrix of order n x q, which relates observations to their random effects

Then consider:

$$var(\boldsymbol{e}) = \boldsymbol{I}\sigma_e^2 = \boldsymbol{R} \tag{2.6}$$

$$var(\boldsymbol{a}) = \boldsymbol{A}\sigma_a^2 = \boldsymbol{G} \tag{2.7}$$

where **I** is a n x n identity matrix, σ_e^2 is the residual error variance, σ_a^2 is the additive genetic variance and $\boldsymbol{A}$ is the relationship matrix. This relationship matrix captures the genetic relationships among the individuals in the pedigree. This matrix is symmetric and the diagonal element for animal *i* (a_{ii}) is equal to 1+F_i, with F_i the inbreeding coefficient of animal *i*. The off-diagonal element, a_{ij}, equals the coefficient of relationship between animals *i* and *j* (Mrode, 1996).

Y, **X** and **Z** are observed, while $\boldsymbol{\beta}$, $\boldsymbol{a}$, **R** and **G** are generally unknown. Solving mixed models involves solving two complementary estimation problems. First, the covariance matrices **G** and **R** need to be estimated. Two methods are widely used: Restricted Maximum Likelihood (REML) and Markov chain Monte Carlo (MCMC). Both methods will be addressed in the following chapters. Second, the vectors of fixed and random effects, $\boldsymbol{\beta}$ and $\boldsymbol{a}$, need to be estimated and will be called the BLUE and BLUP, respectively (Lynch & Welsh, 1998; Schaeffer, 2000; Mrode, 2005).

The BLUP of $\boldsymbol{a}$ can be calculated as:

$$\widehat{\boldsymbol{a}} = \boldsymbol{GZ'V^{-1}}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \tag{2.8}$$

and the BLUE of $\boldsymbol{\beta}$ equals to:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'V^{-1}X})^{-1}\boldsymbol{X'V^{-1}Y} \tag{2.9}$$

where $\boldsymbol{V} = \boldsymbol{ZGZ'} + \boldsymbol{R}$.

The computation of the inverse of $\boldsymbol{V}$ can be challenging. In 1950, Henderson proposed a different form of equations, called the mixed-model equations (MME), to be able to calculate the previous equations more easily (Schaeffer, 2000; Mrode, 2005):

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \boldsymbol{A}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'}\,\boldsymbol{R}^{-1}\,\boldsymbol{Y} \\ \boldsymbol{Z'R}^{-1}\,\boldsymbol{Y} \end{bmatrix} \tag{2.10}$$

With $\alpha\ = \sigma_e^2/\sigma_a^2$ or $(1-h^2)/h^2$.

The solutions to the MME gives the BLUE of $\boldsymbol{\beta}$ and the BLUP of $\boldsymbol{a}$. The full derivation of the BLUP and BLUE estimates can be found in appendix 2A.

In some cases environmental effects can be added as a random factor to the general mixed linear model (equation 2.5). This way the model can account for common environmental effects, such as animals who were reared in the same conditions or born in the same year. In this case the environmental variance (σ_E^2) can be subdivided in the between-group component (σ_c^2) and the within-group or residual variance (σ_e^2).

This addition of environmental effects to equation 2.5 leads to an extended mixed linear model (Mrode, 1996):

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{Wc} + \boldsymbol{e} \tag{2.11}$$

where $\boldsymbol{c}$ is the vector of q common environmental effects and $\boldsymbol{W}$ the design matrix of order n x q, which relates record to their common environmental effects.

It is assumed that $var(\boldsymbol{c}) = \boldsymbol{I}\sigma_c^2, var(\boldsymbol{e}) = \boldsymbol{I}\sigma_e^2$ and $var(\boldsymbol{a}) = \boldsymbol{A}\sigma_a^2 = \boldsymbol{G}$**.**

The MME for the BLUP of $\boldsymbol{a}$ and $\boldsymbol{c}$ and the BLUE of estimable functions of $\boldsymbol{\beta}$ are calculated in the same way as for the simple animal model, leading to the following MME (Mrode, 1996):

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{a}} \\ \widehat{\boldsymbol{c}} \end{bmatrix} \begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} & \boldsymbol{X'W} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \boldsymbol{A}^{-1}\alpha_1 & \boldsymbol{Z'W} \\ \boldsymbol{W'Z} & \boldsymbol{W'Z} & \boldsymbol{W'W} + \boldsymbol{I}\alpha_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'Y} \\ \boldsymbol{Z'Y} \\ \boldsymbol{W'Y} \end{bmatrix} \tag{2.12}$$

where $\alpha_1\ = \sigma_e^2/\sigma_a^2$and $\alpha_2\ = \sigma_e^2/\sigma_c^2$

2.3.4 The accuracy value of an EPD

An important characteristic of an EPD calculation is the accuracy of the calculation (Mrode, 1996). These accuracy values are expressed as a percentage and indicate the quantity of used information. The more information used, the higher the accuracy value. For example, animals without offspring (or offspring without measurement records), will have a lower accuracy value than animals that have offspring with measurements. The value can be interpreted as the correlation between the true and predicted breeding values and indicates the likelihood that an animals EPDs will change over time when more information becomes available. Animals with a high accuracy for a certain EPD will have a true breeding value close to the estimate. Animals with a lower accuracy for a certain EPD can have a true breeding value lower or higher than the estimate (Beef Improvement Federation, 2010; Buchanan & Hanna, 2014; Breedplan, 2015).

The accuracy (*r*) of the predictions is given by the following formula:

$$r_i = \sqrt{1 - \frac{d_i \sigma_e^2}{\sigma_a^2}} \tag{2.13}$$

where r_i is the accuracy value for animal i, d_i is the i^th^ diagonal element of the inverse of $\boldsymbol{Z}'\boldsymbol{Z} + \boldsymbol{A}^{-1}\alpha$ and $d_i\sigma_e^2$ is also called the PEV. The full derivation of equation 2.13 can be found in appendix 2B.

For an EPD, calculated based on the parents' EPDs, the accuracy value is equal to $\frac{1}{2}\sqrt{(r_s^2 + r_d^2)}$, where r_s and r_d are the respective accuracy values for the EPDs of sire and dam (Mrode, 1996).

2.3.5 Restricted maximum likelihood

REML can be used to estimate (co)variances and can therefore be used to estimate the **G** and **R** matrices. Patterson and Thompson developed this method in 1971 and it is based on the maximum likelihood estimation as proposed by Fisher in 1922 (Schaeffer, 2010). These maximum likelihood estimators are the combination of the means (µ) and the variances (σ^2) for every parameter, which fits the observations most likely. The procedure requires $\boldsymbol{Y}$ to have a multivariate Normal distribution. Fishers likelihood function (L) is given by (Schaeffer, 2010):

$$L(\boldsymbol{Y}) = \frac{\exp\left(-0.5(\mathbf{Y}-\mathbf{X\beta})'\boldsymbol{V}^{-1}\,(\boldsymbol{Y}-\boldsymbol{X\beta})\right)}{(2\pi)^{\frac{n}{2}}\sqrt{|\boldsymbol{V}|}}, \tag{2.14}$$

where n is the number of observations. The next step is to maximize the likelihood function by taking the derivative of equation 2.14. This leads to the maximum likelihood function:

$$\ln\left(L(\boldsymbol{Y})\right) = -\frac{n}{2}\ln(2\pi) - \frac{ln|V|}{2} - 0.5(\mathbf{Y}-\mathbf{X\beta})'\boldsymbol{V}^{-1}\,(\boldsymbol{Y}-\boldsymbol{X\beta}) \tag{2.15}$$

REML uses this maximum likelihood function and will correct for the fixed effects in the model by using their estimates, obtained by the least squares procedure. One of the REML algorithms is called expectation maximization REML (EM-REML) and works iteratively. It will use the previous variance estimates, recalculate the $\boldsymbol{V}$ matrix, recalculate the MME and then obtain the new estimates of the variances. This will be repeated until the change in new versus previous estimates is smaller than a pre-specified number (Schaeffer, 2010). Another algorithm is called average information REML (AI-REML). It averages the observed and expected information matrices. This way, inversion of these matrices is much easier. AI-REML has been found to have less problems with correlated components, converges faster than EM-REML, often locates higher maxima of the likelihood function and performs well in animal breeding context (Gilmour *et al.*, 1995).

2.3.6 The Markov chain Monte Carlo method

The MCMC method is based on Bayesian statistics. Bayesian statistics uses probabilities to derive a joint posterior distribution from the given data. This joint posterior distribution is a probability distribution associating each value of a parameter to a probability. Bayesian statistics is based on the Bayes' theorem (Olshausen, 2004):

$$p(\boldsymbol{\theta}|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\boldsymbol{Y})} \tag{2.16}$$

where $\boldsymbol{\theta}$ is a vector of parameter variables and $\boldsymbol{Y}$ is the given data vector. The term $p(\boldsymbol{Y}|\boldsymbol{\theta})$ is the probability of the observed data $\boldsymbol{Y}$, given the parameters $\boldsymbol{\theta}$, called the likelihood function. The term $p(\boldsymbol{\theta})$ is called the prior probability and reflects the prior knowledge. The specification of the prior is often the most subjective aspect of Bayesian statistics. The term $p(\boldsymbol{Y})$ is obtained by integrating $p(\boldsymbol{Y}|\boldsymbol{\theta})\, p(\boldsymbol{\theta})$ over $\boldsymbol{\theta}$ and plays the role of a normalizing constant. Finally, the term $p(\boldsymbol{\theta}|\boldsymbol{Y})$ is known as the posterior probability and reflects the probability of the vector of random variables after consideration of the data (Kass *et al.*, 1997; Olshausen, 2004; de Villemereuil, 2012;).

The MCMC method works by constructing a Markov chain. A Markov chain is a stochastic process which samples random values independent from the history of the chain. This sampling technique is called Monte Carlo. These two principles together form the MCMC method. MCMC will not systematically go through every possible combination of parameter values, but moves stochastically through the parameter space to calculate probabilities for every parameter, such as the means, the variances and correlations. The result of the MCMC method will be a joint posterior distribution (Kass *et al.*, 1997; de Villemereuil, 2012; Hadfield, 2016). This distribution can be characterized by the highest posterior density interval. This highest posterior density interval is the narrowest credible interval that indicates the sample space that covers a given probability (Plummer *et al.*, 2006). In unimodal posterior distributions, the mean or mode can be used as estimator (Hadfield, 2016).

MCMC can be used to construct the posterior distribution in animal models for the **R** and **G** matrices and also for all levels of the fixed and random effects.

A detailed derivation of the posterior distributions for the parameters in the animal model is given in appendix 3.

2.3.6.1 *Sampling algorithms*

One of the most used sampling algorithms is called Gibbs sampling (Hadfield, 2016). This algorithm obtains observations of a sequence that will approximate the posterior probability distribution by picking for every parameter in the distribution random coordinates from a multivariate Normal distribution. These values are added to the list of previous coordinates. The algorithm will use an accept-reject step to decide whether it will move to these next values. If the posterior probability for this new set of parameter values is greater than the posterior probability of the current set of parameters, the algorithm will ‘move' to these new values. If this is not the case, the chain will move there only occasionally, depending on the relative difference between the old and new posterior probabilities.

2.3.6.2 *The chain length*

When running a sampling algorithm, the model iterates a pre-defined number of runs. The length of the Markov chain will depend on the chosen computational power and time. In general, the longer the chain, the better, since it will increase the effective sample size. This effective sample size is defined as the number of uncorrelated, independent samples in the Markov chain. de Villemereuil (2012) recommends an effective sample size of at least 1000 samples.

A tool to check the behavior of the chain is plotting the trace of the values. This is a visualization of the evolution of the sampled values along the iterations. This way the convergence of the trace can be checked and it should be verified that the chain does not get ‘stuck’ on a certain value. Apart from the trace plot, it is useful to plot the posterior density function for each parameter.

2.3.6.3 *The burn in period*

Depending on the starting values, the initial values of the Markov chain will often not be representative to the values further in the chain. It may take a number of samples before the chain reaches an ‘equilibrium’ state. It is important that the values from this phase of converging are not mixed up with the values that follow. The period from start to the point that convergence has happened is called the burn-in period. In figure 2.2, an example of a trace is plotted where the burn-in period is clearly visible. Unfortunately, it is not possible to predict the burn-in period in advance, so post hoc checks are necessary (Kass *et al.*, 1997; de Villemereuil, 2012; Hadfield, 2016).

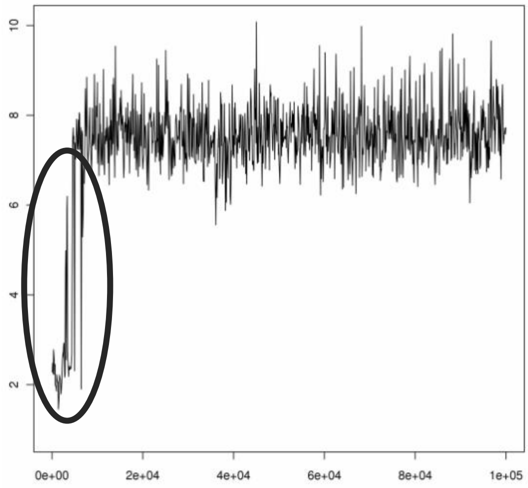

Figure 2.2 An example trace from Hadfield (2016) where the burn-in period is indicated. After the burn-in period, the Markov chain converges to a value of ± 7.5.

2.3.6.4 The autocorrelation of the chain and the thinning interval

Because of the sampling procedure, successive iterations will be correlated to each other. This is caused by the fact that the previous value of the other parameters determines whether or not the new value will be accepted. This tendency is called autocorrelation and reduces the effective sample size of the chain. A way of dealing with this autocorrelation is by using a thinning interval. This thinning interval will only keep one in every *n* iterations. It will reduce autocorrelation and saves memory space (Schaeffer, 2000; de Villemereuil, 2012; Hadfield, 2016).

2.3.6.5 Prior distributions

In order to determine the posterior distribution of **Y**, there is need for a prior probability distribution for the parameters. Most of the time, a weak non-informative prior is used. This means that the influence of the prior on the estimated posterior distribution will be minimal (de Villemereuil, 2012). When the effective sample size is sufficiently long, the influence of the prior becomes negligible (Schaeffer, 2000; de Villemereuil, 2012).

A frequently used prior for fixed effects is a uniform distribution or a wide Normal distribution (de Villemereuil, 2012; Hadfield, 2016). One of the priors commonly used for the residual and random effect variances is the inverse-Wishart distribution. This distribution has two parameters: *V* and *nu* and tends to a point mass on *V* and is right skewed when *nu* is small. The inverse-gamma function is the univariate derivate of this inverse-Wishart function (Hadfield, 2016). The probability density function of an inverse-gamma function with a variance limit of 1 (*V*=1) and a varying degree of belief parameter (*nu*) can be found in figure 2.3. The inverse-gamma function is in its turn a special case of the scaled inverse chi-squared distribution.

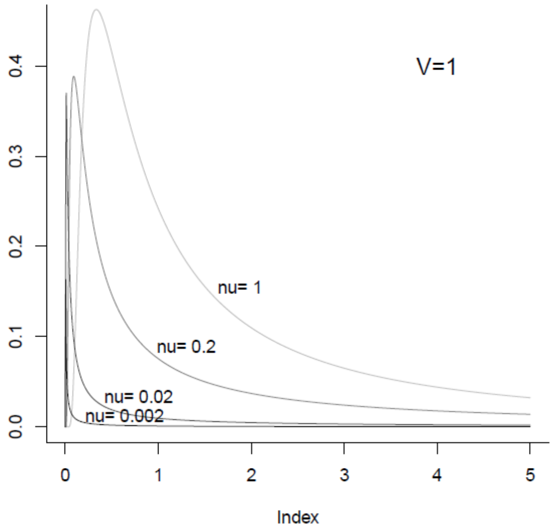

Figure 2.3 The probability density function of an inverse-gamma function with V=1 and nu varying from 0.002 to 1
(figure from Hadfield, 2016)

2.3.7 The differences between MCMC and REML

de Villemereuil *et al.* (2013) studied the differences between REML and MCMC in an animal model. They found that the animal model using the REML algorithm had a slight advantage in estimating traits that were normally distributed and the MCMC method had slightly better estimates for categorical traits. The heritability estimate was more accurate when estimated by REML. They concluded that the main advantages of the MCMC method are based on the possibility of fitting a great variety of non-Normal distributed data. Hadfield (2016) concluded that REML is faster and more easy to use, while MCMC can be slow and more difficult. The definition of the prior distribution can be challenging (de Villemereuil *et al.*, 2013; Hadfield, 2016). The accuracy of the approximation of MCMC increases the longer the analysis is run and can be more exact than the REML method (de Villemereuil *et al.*, 2013; Hadfield, 2016).

As an alternative for MCMC, another Bayesian method is being developed: the integrated nested Laplace approximation. This method looks promising and less computation intensive but still needs further development (de Villemereuil *et al.*, 2013; Maniatis *et al.*, 2015).

2.4 The phenotypic and genetic correlation of traits

When analyzing different traits on the same individual, some traits can show a phenotypic correlation. Such phenotypic correlation can have two causes. The first is called the environmental correlation. This correlation is caused by animals sharing the same environment (e.g. the same rearing conditions or the same dam). The second type is called the genetic correlation and is most of the time caused by genes influencing multiple traits at the same time, a phenomenon known as pleiotropy (Lynch & Walsh, 1998; Nicholas, 2010; Acquaah, 2012).

Correlations are either positive or negative and always range from -1 to +1. A correlation of zero means that traits are not (linearly) related. A positive correlation means that selection for one trait leads to an increase in the other trait, for example weaning weight and yearling weight in cattle (Bourdon & Brinks, 1986). A negative correlation indicates that selection for one trait leads to a decrease in the other trait (e.g. the negative correlation between birth weight and calving ease) (Nicholas, 2010). Positive and negative is not equal to favorable or unfavorable as this depends on the measurement scale of the variables.

Genetic correlations create the possibility of indirect selection. Therefore, it is possible to select for one trait and at the same time influencing the other trait in a desired direction. This indirect selection can become very useful when breeders want to improve traits which are harder or more expensive to measure. Breeders can improve the desired trait, while only measuring the genetically correlated trait (Nicholas, 2010).

The correlated response (CR) can be calculated as following (Acquaah, 2012):

$$CR = \Delta G_{2(1)} = I * h_1 * r_{g_{12}} * \sigma_{a2} \tag{2.17}$$

where trait 1 is the measured trait, trait 2 the trait to improve, $\Delta G_{2(1)}$ the genetic progress of trait 2 by selection for trait 1, *I* the selection intensity, h_1 the reliability of the heritability estimate of trait 1 (which is equal to $\sqrt{h_1^2}$), $r_{g_{12}}$ the genetic correlation between trait 1 and trait 2 and σ_{a2} the genetic standard deviation of trait 2. The efficiency of indirect selection is calculated as (Acquaah, 2012):

$$\frac{\Delta G_{2(1)}}{\Delta G_2} = \left| r_{g_{12}} \frac{h_1}{h_2} \right| \tag{2.18}$$

where ΔG_2 is the genetic progress in case of direct selection for trait 2 and h_2 the reliability of the heritability estimate of trait 2 (which is equal to $\sqrt{h_2^2}$).

2.5 Crossbreeding

Crossbreeding is defined as the mating of individuals from two different populations or breeds. In the USA, it is a common practice in the beef industry. Animals resulting from a crossing are called crossbreds. Crossbreeding is, together with selection, an important tool to change the genetic composition of a herd or breed. The crossing of two different cattle breeds leads to a genetic enrichment of the offspring.

There are two main advantages of crossbreeding (Nicholas, 2010). First, a crossbred individual often shows a higher performance than expected based on the means of the two parent populations. This higher performance is called heterosis or hybrid vigor. Heterosis is defined as the observed difference between the performance of the offspring and the mean performance of the two parent populations. It is caused by dominant and epistatic gene effects and the difference in allele frequencies between the two populations. The higher the genetic diversity between the two breeds, the higher the heterosis of the crossbred. Heterosis is mostly observed in traits with a low heritability, thus the traits that improve slow by selection. Traits with a high heritability are less affected by heterosis (Nicholas, 2010).

A second benefit of crossbreeding is using the complementarity of two breeds. Breed complementarity is the combination of the strengths of both breeds in the cross. These strengths of the breed are not only performance traits but can also concern environmental adaptation or coat color (Nicholas, 2010).

3 The Black Hereford breed and its origin

The predominant beef cattle breed in the USA is Angus. In 2016 the American Angus Association registered more than 334000 animals (American Angus Association, 2016a, 2016b). Estimations say that over 60 % of all cattle raised in the USA have Angus blood (American Angus Association, 2016b). The second largest cattle breed in the USA is the Hereford, with 79082 registries in 2016 at the American Hereford Association (American Hereford Association, 2016). The top five breed list is completed by Gelbvieh, Charolais and Simmental (American Angus Association, 2016b). These beef cattle breeds are *Bos taurus* cattle (Felius, 2016). *Bos indicus* cattle breeds are a minority in the USA (Felius, 2016).

3.1 The Angus breed

Figure 3.1 An Angus bull (Watercolor image by Marleen Felius (1995))

The Angus breed was founded in Aberdeen, Scotland and brought to the USA in 1873 (Felius, 2016). The cattle is black coated and known for producing well-marbled meat of prime quality (American Angus Association, 2016b). They are naturally polled, calve easily, have a low birth weight and are fast growing (Felius, 1995). Mature cows weigh typically 650 kg and bulls weigh up to 1000 kg (Felius, 1995). Some disadvantages of Angus cattle are their higher feed conversion in comparison to other beef breeds and the fact that Angus bulls can be aggressive and dangerous to handle (American Angus Association, 2016b). Although most Angus are black, there is a recessive red allele present in the population (Ekarius, 2008; Phillips, 2010). The red Angus population is considered as a different breed and is registered at the Red Angus Association. Blacks are registered at the American Angus Association (American Angus Association, 2016b).

In order to promote the Angus breed, the American Angus Association created in 1978 the Certified Angus Beef® program (American Angus Association, 2016b). The program markets premium Angus beef to restaurants and retailers. The Certified Angus Beef® label is recognized in 53 countries and marketed over 460 million kg of beef in 2016 (American Angus Association, 2016b). This premium beef label increased the demand for Angus cattle (American Angus Association, 2016b; Hoagland, 2016). Before a carcass can be sold as Certified Angus Beef®, it is screened for the right specifications. To receive the label, cattle must be at least 51% black-coated and meet ten different criteria concerning beef quality (e.g. marbling, fat thickness and rib area size) (American Angus Association, 2016b).

The demand for black coated cattle increased dramatically due to the requirements for this label. Black calves are worth more than red ones, as much as 50 to 100 $ per calf (American Angus Association, 2016b). This stimulated some breeds to develop a black strain, for instance the Black Limousin, Black Chianina, Black Simmental and the Black Hereford (Felius, 2016; Hoagland, 2016).

3.2 The Hereford breed

Figure 3.2 A Hereford bull (Watercolor image by Marleen Felius, (1995))

The Hereford breed originates from Herefordshire County in Western England. In the beginning, the breed was used for milk, meat and draught purposes, and switched later to a beef breed (Felius, 1995). It is easily recognizable because of its red body and white face, belly, feet and tail switch. The typical white face phenotype is called 'baldie' and is a dominant trait (Franke, Burns, & Koger, 1975; Grosz & MacNeil, 1999; Schmutz, 2016). The baldie face in crossbreds makes clear there is a Hereford involved in the cross. Herefords have a good fertility, fatten easily and are relative docile (Felius, 1995). They can withstand both fairly low and high outside temperatures. The breed performs good in grass-based production as well as in the feedlot. Mature cows weigh 600-800 kg and bulls weigh 1000-1200 kg (Felius, 1995). Herefords are naturally horned but in the 1890s breeders developed a polled strain, called the Polled Hereford. The Hereford was the first British cattle breed imported in the USA in 1839 (Ekarius, 2008; Phillips, 2010; Felius, 2016). Since the 1980s, the Hereford has come under competition from the other beef breeds and went from 253832 new registered calves in 1970 to 97424 in 1990 (Felius, 1995).

3.3 Commercial cattle crossbreeding

Commercial cow-calf producers often use breed complementarity to produce their calves. In many commercial herds, the main proportion of the herd is Angus-based, black cattle (American Angus Association, 2016b; Hoagland, 2016). These black cows are crossed with well-performing beef bulls. In many cases a Hereford bull is used. This way, the efficiency in the above mentioned traits of the Hereford is combined with the marbling and meat quality of Angus (Felius, 1995; Hoagland, 2016; American Angus Association, 2016b). Long (2009) found in his study that the use of Hereford bulls as sires on an Angus cow herd led to better performing offspring in comparison to Angus-sired calves. Hereford-sired cattle were more cost efficient in comparison to the Angus-sired offspring. In the end, the Hereford-sired calves brought more money, although they had a lower acceptance rate for the Certified Angus Beef® program (20 % instead of 30 % for the Angus-sired cattle) (Long, 2009). Other studies indicated that Hereford x Angus crossbred calves had economic advantages over purebred Angus calves (Denton, 2009a, 2009b; Gugelmeyer, 2009)

One major disadvantage of the common Angus x Hereford cross is the risk of producing red coated calves. These calves will not be allowed to register for the Certified Angus Beef® program.

3.4 The Black Hereford breed

Figure 3.3 A Black Hereford bull (own picture)

One way to reduce the risk of producing a red coated calf in a commercial crossbreeding system is by crossing the Angus herd to a black bull. This led to the development of the Black Hereford breed. The Black Hereford arose from a population of Herefords by crossing these with Angus cows. The Black Hereford still looks fairly the same as the regular Hereford and has the same major characteristics except for its black coat. The main purpose of the Black Hereford breed is to deliver sires that can be used to be bred to a black-coated cow. By using this cross, cow-calf producers are able to use Hereford genetics and thus breed complementarity, while at the same time reducing the risk of a red (discounted) calf (Hoagland, 2016).

The Black Hereford breed is a docile breed with good meat quality and a good feed efficiency (Hoagland, 2016). Most of the cattle is naturally polled and cows have good mothering abilities. Mature males weigh between 900 and 1000 kg, while mature females weigh on average 600 kg (Hoagland, 2016).

The initial crossings to create the Black Hereford breed date back to the mid-1990s (Hoagland, 2016). Selected Hereford bulls were mated to Angus cows. The offspring of this cross (called F_1) are almost all black baldie calves. These calves are 50% Hereford and 50% Angus. The F1-females were bred back to a set of Hereford bulls to increase the percentage of Hereford (Hoagland, 2016). This type of cross is often called a grading up cross. It took a few generations before a real breed was established. In the following figure (3.4), the crossing scheme is illustrated. The first registered Black Hereford bull, JN Balder 7504, was born in 1997. Since 2004 is the Black Hereford breed one of the fastest growing cattle breeds in the USA. Now (2017), there are approximately 150 active Black Hereford breeders spread over 30 states in the USA (Hoagland, 2016).

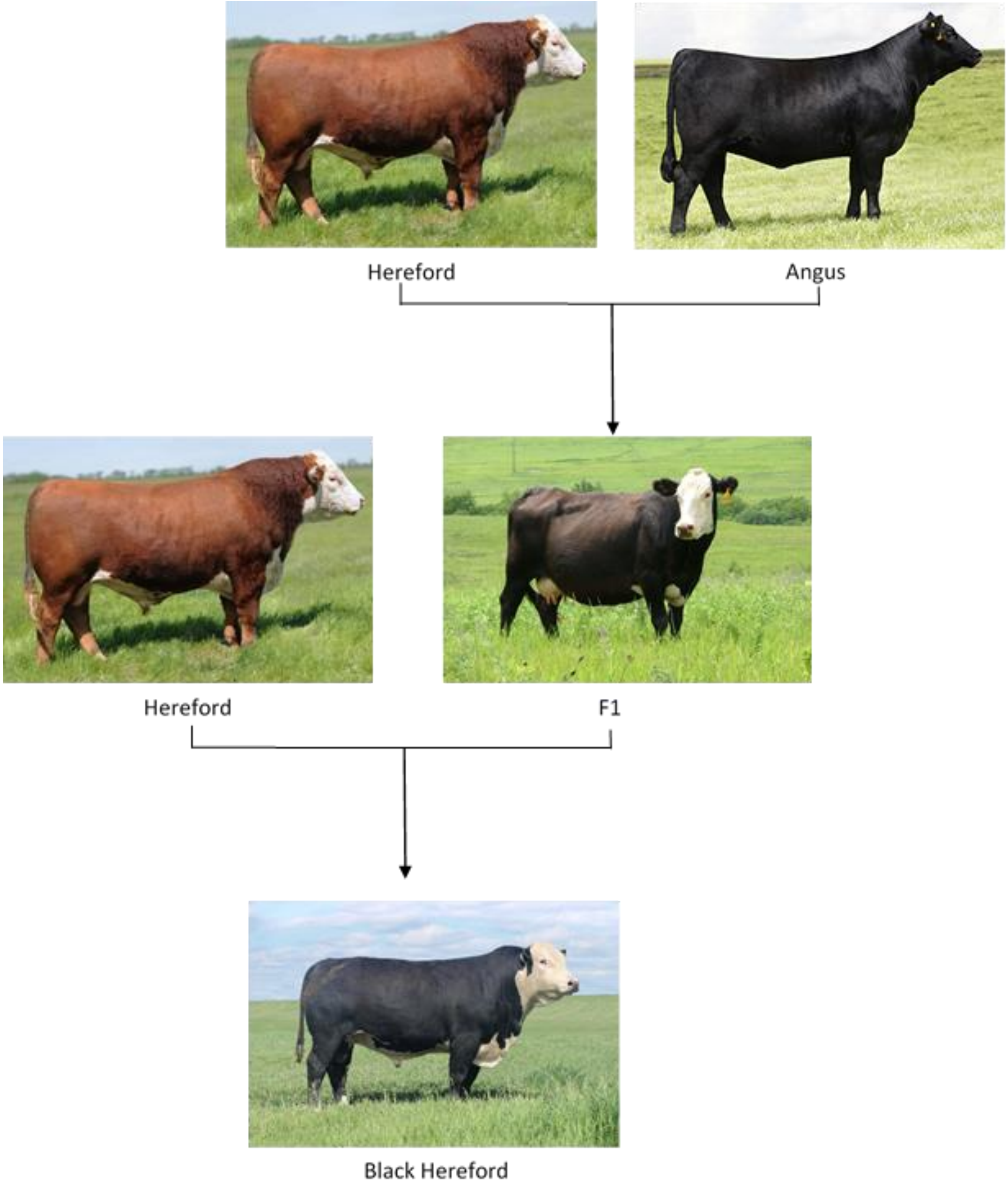

Figure 3.4 The crossing scheme that has been used to create the Black Hereford breed

The breeding goal of the Black Hereford population over the last few years was to consolidate the Hereford genetics in the breed. The main focus at the reproduction side lies on increasing the ease of calving, producing low birthweight calves and having heifers calf for the first time at an age of two years old (Hoagland, 2016). The focus at the production side lies on better marbling of the meat and increasing the ribeye area, while maintaining the good feed efficiency. Besides these focusses, attention is given to the eradication of the last horned alleles present in the population (Hoagland, 2016).

To register an animal as a Black Hereford at the American Black Hereford Association (ABHA), it has to have at least 62,5 % (5/8th) of its genetics from Hereford, the other 37,5 % or less being Angus genetics (Hoagland, 2016). A purebred Black Hereford needs to be at least 93 % Hereford based. All registered Black Herefords must have a black coat but they do not have to be homozygous black. Most of them have the typical white face, abdomen and feet of a Hereford (Hoagland, 2016).

The Black Hereford breed must not be confused with the European Black Hereford crossbred between Hereford bulls and Holstein-Friesian dairy cows. This cross has the purpose of producing calves suitable for beef production and starting a new lactation cycle for the dairy cow. Neither should the Black Hereford be confused with the Australian term 'Black Baldie' which is a collection of black baldie cattle.

The first EPDs for the Black Hereford breed were calculated in 2005 by Dr. Dan Moser. In 2017, there are EPD calculations available for five traits which are calculated yearly: birth weight, weaning weight, yearling weight, maternal milk production and total maternal traits (Hoagland, 2016). Some breeders also report EPDs for other traits like the maternal milk and growth EPD.

3.5 The black coat color

The coat color of cattle is determined by several genes but the most important locus is the Extension (E) locus with three possible alleles: ED, dominant black, E+, the wild-type allele responsible for most combinations of red or reddish brown and black, and e, recessive red (Olson, 1999; Schmutz, 2016). Black Angus cattle has at least one ED allele, and Herefords are homozygous e. Angus cattle can carry an E+ or e allele (Nicholas, 2010; Olson, 1999; Schmutz, 2016). When Hereford and Angus are crossed, calves can be black or red coated. A heterozygous black Angus cow has 50 % chance of producing a red calf, when sired to a regular Hereford bull. When that same Angus cow gets sired by a heterozygous Black Hereford bull, this chance reduces to 25 %. If the Angus cow is sired to a homozygous Black Hereford bull, the calf will always be black. A homozygous black Angus cow will always produce a black calf, sired either by a regular Hereford or Black Hereford bull. The main problem of heterozygous black cows is that they are not distinguishable from homozygous black cows by eye. This can only be done by performing a DNA-test. In commercial conditions, this DNA-test is too expensive to perform on large scale on the whole herd (Hoagland, 2016).

4 The reproduction of cattle

4.1 The puberty of cattle

In mammals, puberty is induced by changes in hormone levels of luteinizing hormone (LH) and follicle stimulating hormone (FSH) (Bouron & Boulpaep, 2003). LH and FSH are both secreted by the anterior pituitary gland in the brain. They stimulate the gonads to increase activity. In males, LH increases the testosterone production (Bouron & Boulpaep, 2003). Testosterone will increase sperm production, stimulate accessory sex gland growth, increase sexual behavior and produce male secondary sex characteristics. FSH secretion will stimulate Sertoli cell function in the testes (Bouron & Boulpaep, 2003). In females, FSH and LH allow follicles to develop and produce estrogen (Bouron & Boulpaep, 2003). Estrogen will cause a positive feedback effect leading to a LH surge. This LH surge causes ovulation (Bouron & Boulpaep, 2003). The start of these hormonal expressions indicate in both males and females the onset of puberty.

Land (1973) was the first to discover a link between the development of male and female reproductive traits in sheep and mice. He found that hormonal expression (e.g. LH and FSH) and sexual development was equal in both sexes. The same relationship can be found in cattle (Toelle & Robison, 1985; Moser *et al.*, 1996).

The age at puberty (AP) of a heifer is defined as the age at which the first ovulatory estrus is detected and thus indicates the onset of puberty (Martinez-Velázquez *et al.,* 2003; Wathes *et al.*, 2014). AP is used as a measure of heifer fertility and is an important reproduction trait in both dairy and beef cattle production systems. Although puberty is detectable in research settings, it is hard to observe in field populations, since it needs blood sampling and subsequent laboratory analyses (Wathes *et al.*, 2014). The AP of heifers depends on environmental conditions and feed. Inadequate nutrition will delay puberty (Cammack *et al.*,2009).

In commercial breeding conditions heifers need to cycle a couple of times before being exposed to breeding (Wathes *et al.*, 2014). This way, they will be physiologically more ready for conception and gestation. The length of the bovine estrus cycle is 21 days and thus, puberty should occur at least 6 weeks before the target breeding date (Perry, 2004; Wathes *et al.*, 2014).

The AP of heifers is breed-dependent. Laster *et al.* (1972) found that purebred Hereford heifers have an average AP at 389.5 ± 12.9 days, while purebred Angus heifers had an AP around 372.2 ± 10.0 days. The AP in Black Hereford has never been studied. Patterson *et al.* (1992) found that AP is related to both a certain age and a certain weight. They also indicated the importance of preweaning and postweaning nutrition levels. It is hard to conclude whether age or body weight is most determinant for AP (Nelsen, Long, & Cartwright, 1982; Patterson *et al.*, 1992).

4.2 The gestation and the optimal age of calving

The gestation period of cattle is on average 283 days (Casas *et al.*, 2012; Hoagland, 2016). Heifers or smaller cows may calve a few days earlier. Gestation in Angus cows can be up to four days shorter than in Hereford cows. The main cause of this gestation difference is that Angus has been intensively selected for lower birth weight, which coincided with calving a few days earlier (Casas *et al.*, 2012; American Angus Association, 2016b; Hoagland, 2016).

In the USA, most breeders want to have their beef heifers calving for the first time at an age of 2 years. From then on they will calve with a one-year interval until an age of 6 to 10+ years. This means they have to be bred when they are about 15 months old. 15 months is widely accepted as the minimum age at which beef heifers are sufficiently grown to be exposed to breeding (Wathes *et al.*, 2014). A calving age of 2 years is necessary to maintain the seasonality of the production cycle and calving pattern. It reduces the non-productive period of the heifer without compromising the health and longevity (Wathes *et al.*, 2014). Herefords reach puberty on average at an age of 14 to 15 months old, Angus between 13 and 14 months old (Morris, Baker, & Cullen, 1992; Smith, Brinks, & Richardson, 1989a).

Day and Nogueira (2013) found that heifers, that calve for the first time at an age of 2 years, produce on average 0.7 more calves than heifers, calving first at 3 years of age, by the time the cows were 6.5 years of age. An economic analysis indicated that the difference in profit was 500 $/cow for a cow with an age at first calving (AFC) of 2 years instead of 3 years at the end of 4 productive years. Although the first calves from 2 year old heifers were lighter at weaning, no differences were found in the following calves. Nunez-Dominquez *et al.* (1985) found an AFC of 2 years had no adverse consequences on subsequent reproduction and maternal performance. Cows with an AFC of 2 years had a larger chance of producing an extra calf in their lifetime over cows with an AFC of 3 years. They also produced on average 138 kg more of weaned calf weight and their economic efficiency was 6 % to 8 % higher than cows bred to have their first calf at an age of 3 years. They also found that cows with an AFC of 2 years and a cumulative lifetime production up to 12 years of age made about 2000$ more income than cows with an AFC of 3 years.

Even within the group of heifers that calve at approximately 2 years of age there is a difference in lifetime production between the heifers who calve early and late in the calving season (Hoagland, 2016). An earlier calving results in more and better lactation, mostly because of a more optimal forage availability during the lactation peak. They also have more days postpartum at the start of the next breeding season and cows will have more time to recover. They will be more likely to have resumed normal estrous cycles and fertility by then. Heifers that calve early in the calving season will generally stay one year longer in the herd, resulting in one more calving during their lifespan (Patterson *et al.*, 1992; Hoagland, 2016).

5 Scrotal circumference

5.1 Scrotal circumference

The scrotal circumference (SC) of a bull is defined as the distance around both testes and is expressed in centimeters (cm). The SC measurement is carried out with a circular ‘scrotal’ tape and is part of the breeding soundness evaluation performed on breeding bulls (see figure 5.1). The SC is taken at the largest diameter of the scrotum (Beef Improvement Federation, 2010). The measurement is relatively simple and inexpensive to perform. (Lunstra *et al.*, 1988; Martinez-Velázquez *et al.*, 2003; Hoagland, 2016). When the SC measurement is performed at yearling age, it is called the yearling scrotal circumference (YSC).

Figure 5.1 Left: A scrotal tape, used for the scrotal circumference measurement. (own picture)
Right: The use of a scrotal tape in practice on a yearling Black Hereford bull (picture: Joe Hoagland)

In practice, it is almost impossible to measure every bull when he is exactly one year old. Instead, the whole herd is measured at the same moment and an individual adjustment for age is made (Bourdon & Brinks, 1986; Lunstra *et al.*, 1988). An adjustment by age is found to be more accurate than an adjustment by body weight at measurement (Bourdon & Brinks, 1986; Lunstra *et al.*, 1988; Beef Improvement Federation, 2010). The age adjustment factor is breed specific. Lunstra *et al.* (1988) and the Beef Improvement Federation (2010) both suggest that, besides age adjustment, adjustment for the age of the dam is also applicable. In table 5.1 and 5.2 some adjustment factors for age and age of dam can be found for Angus and Hereford cattle. There are currently no adjustment factors estimated for the Black Hereford breed. The adjustment formula that can be used is (Lunstra *et al.*, 1988; Beef Improvement Federation, 2010):

$$Adj_{365d}\,YSC = YSC + f * (365 - age) + AD \tag{5.1}$$

where f is the age adjustment factor and AD the age of dam adjustment factor.

Table 5.1 Age adjustment factors (f) for YSC in yearling bulls between 300 and 400 days of age by different authors for Hereford and Angus cattle

Author	Year	SC adjustment factor (cm/day)		
		Angus	Hereford	Polled Hereford
Bourdon and Brinks	1986		+ 0.0260	
Lunstra *et al.*	1988	+ 0.0340	+ 0.0360	
Geske *et al.*	1995	+ 0.0374	+ 0.0425	+ 0.0305

Table 5.2 Scrotal circumference (SC) adjustment factors for age of dam (Lunstra *et al.*, 1988)

Age of dam (years)	Adjustment factor for YSC (cm)
2	0
3	- 0.4
4	- 0.8
> 4	- 1.3

The YSC adjustment factors for age of dam in table 5.2 adjust the observed values to a dam of 2 years old. In other words, the YSC of bulls from 3 year old dams is on average 0.4 cm larger, those of 4 year old dams 0.8 cm and those of dams older than 4 years 1.3 cm in comparison to a bull of a one-year old dam.

5.2 Yearling scrotal circumference measurements and heritability estimates

A list of averages of YSC measurements from previous studies can be found in table 5.3. This table also contains heritability estimates for YSC and/or age-adjusted YSCs. In all but one studies, YSC has been found to be a moderate to highly (>0.35) heritable trait. These heritabilities suggest that selection for YSC is possible.

Table 5.3 Overview of yearling scrotal circumference (YSC) measurements in different studies on different beef breeds. The mean YSC (in cm) and the standard deviation of YSC (in cm) are given together with the heritability (h^2) estimate for YSC or age-adjusted YSC (Adj-YSC) (adjusted to an age of 365 days)

Author	Year	No. of bulls	Breed	Method used	YSC Mean (cm)	YSC SD (cm)	h^2 YSC	h^2 Adj-YSC
Neely et al.	1982	401	Hereford	PHS	31.1	2.5	0.44 ± 0.24	0.44 ± 0.24
Knights et al.	1984	717	Angus	REML	35.7	2.1	0.36 ± 0.06	-
Toelle and Robison	1985	528	Hereford	PHS	31.3	2.6	-	-
Bourdon and Brinks	1986	4233	Hereford	PHS	34.4	2.1	0.53 ± 0.06	0.49 ± 0.06
Lunstra et al.	1988	3090	Total of 12 breeds	PHS	32.3	-	-	0.41 ± 0.06[a]
Kriese et al	1990	10511	Hereford	SGM	34.3	2.7	-	0.53
Moser et al.	1996	407	Limousin	REML	32.7	1.3	-	-
Evans et al	1999	1220	Hereford	REML	31.1	-	-	0.71
Arthur et al.	2001	7260	Angus	REML	35.2	2.9	0.43 ± 0.06	-
Martinez-Velázquez et al.	2003	7580	Total of 12 breeds	REML	32.6	3.0	0.41 ± 0.04	-
Kealey et al.	2006	626	Hereford	REML	35.0	2.1	0.57 ± 0.09	-
McAllister et al.	2011	43487	Red Angus	REML	35.2	2.7	0.32 ± 0.09	-

With PHS = paternal half-sib analysis, REML = restricted maximum likelihood, SGM = sire-maternal grandsire model
Values indicated by a (-) were not reported or included in the study
[a] Adjusted to an age of 354 days

5.3 Scrotal circumference and production traits

Bourdon and Brinks (1986) found in Hereford bulls that age-adjusted YSC was positively genetic correlated with adjusted birth weight, weaning weight and yearling weight (0.18, 0.29 and 0.44, respectively). Adjusted YSC was also positively genetic correlated with postweaning average daily gain (0.35). Neely *et al.* (1982) found genetic correlations of 0.86, 0.52 and 0.22 between adjusted YSC respectively with adjusted weaning weight, yearling weight and postweaning gain for a set of 401 Hereford bulls. Knights *et al.* (1984) found a genetic correlation of 0.68 between YSC and yearling weight in Angus bulls. These findings were in line with the results of Kriese *et al.* (1991), who found a moderate positive additive genetic correlation between adjusted YSC and adjusted postweaning gain These studies indicate that there is a favorable genetic relationship between YSC, growth and growth rate in bulls.

Part of this relationship can be explained by the fact that bulls that grow faster, have also grown a larger scrotum at yearling age (Bourdon & Brinks, 1986). Because of the favorable genetic relationship, selection for YSC will not negatively affect these economically important growth traits (Kriese *et al.*, 1991).

Smith *et al.* (1989b) found that for each additional centimeter of YSC of the sire, the weaning weight, yearling weight and average daily gain increased in the offspring.

McAllister *et al.* (2011) found genetic correlation estimates near zero between YSC and intramuscular fat percentage (0.05) and between YSC and marbling score (0.01). Selection for YSC will not influence meat quality traits (McAllister *et al.*, 2012).

5.4 Scrotal circumference and male reproduction traits

Previous research from Hahn *et al.* (1969) found a highly positive correlation (0.81) between YSC and total sperm production per ejaculate in dairy bulls. Brinks *et al.* (1978) found a positive correlation between YSC and semen quality in *Bos taurus* cattle herds (Brinks *et al.*, 1978 cited by Burns (2011)). They also found that by an increase of YSC, there was an increase in sperm motility, percent normal sperm, sperm volume and sperm concentration and a decrease in sperm abnormalities. Similar results were found by Kealey *et al.* (2006). So in general, genetic correlations indicate a positive association between a bulls YSC and his own fertility (Beef Improvement Federation, 2010).

5.5 Scrotal circumference and female reproduction traits

5.5.1 The age at puberty

The age at puberty of heifers is dependent on the breed and management conditions (e.g. feed level) (Wathes *et al.*, 2014). Heritability estimates for age at puberty (AP) can be found in table 5.4. Of all fertility traits in heifers, AP often has the highest heritability (Wathes *et al.*, 2014).

Research has shown a moderate to high negative genetic correlation between a sires YSC and AP in his daughters (Smith *et al.*, 1989b; Morris *et al.*, 1992; Beef Improvement Federation, 2010). Smith *et al.* (1989b) found a regression coefficient of -0.796 days/cm. This means that for every increase of 1 cm of a sires SC, his daughters reach puberty 0.796 days earlier. Ludwig (2012) reported a regression coefficient of -4 days/cm. Moser *et al.* (1996) found that selection for bulls with a larger YSC led to daughters reaching puberty at significantly earlier ages. In table 5.4, several genetic correlation estimates between YSC and AP can be found.

Table 5.4 Several heritability estimates (h^2) for age at puberty (AP) and genetic correlations between a sires yearling scrotal circumference (YSC) and AP in daughters in different beef breeds

Author	Year	Number of heifers	Method used	AP h^2	Genetic correlation between YSC and AP
Brinks *et al.*[a]	1978	-	PHS	-	-0.71
Lunstra [b]	1982	-	PHS	-	-0.98
Smith *et al.* (b)	1989	779	PHS	0.10	-
Morris *et al.*	1992	1302	REML	0.33	-0.41[c]
Splan *et al.*	1998	2864	REML	0.47	-
Martinez -Velázquez *et al.*	2003	5292	REML	0.16	-0.15
Gargantini *et al.*	2005	1184	REML	0.52	-0.57

[a] *cited by Toelle and Robison, 1985 and Burns, 2011*
[b] *cited by Martin et al., 1992*
[b] *YSC adjusted for an age of 334 days*
With PHS = paternal half-sib analysis, REML = restricted maximum likelihood.
Values indicated by a (-) were not reported or included in the study.

5.5.2 The age at first calving

The age at first calving (AFC) is another important trait of heifer fertility. AFC can be calculated from the birth dates of mother and calf. AFC is highly determined by management conditions (e.g. the moment of breeding) and nutrition. Some heritability estimates for AFC can be found in table 5.5.

Table 5.5 Several heritability estimates (h^2) for age at first calving (AFC) and genetic correlations between a sires yearling scrotal circumference (YSC) and AFC in daughters in different beef breeds

Author	Year	Number of heifers	Method used	AFC h^2	Genetic correlation between YSC and AFC
Toelle and Robison	1985	645	PHS	0.23	-0.38
Smith *et al.* (a)	1989	779	PHS	0.01	-
Gutiérrez *et al.*	2002	2533	REML	0.24	-
Martinez - Velázquez *et al.*	2003	4835	REML	0.08	-
Berry and Evans	2014	64380	REML	0.31	-
Berry *et al.*	2014	-	REML	0.14	-

With PHS = paternal half-sib analysis, REML = restricted maximum likelihood.
Values indicated by a (-) were not reported or included in the study.

AFC is highly correlated with age at subsequent calvings and the interval between subsequent calvings (Gutiérrez *et al.*, 2002). Smith *et al.* (1989a) concluded that earlier ages at puberty were associated with earlier ages and dates of calving. Smith *et al.* (1989b) found a regression coefficient for AFC of -0.667 days for every centimeter increase in a sires SC.

These favorable correlations between YSC, AP and AFC indicate that selection for a larger YSC can reduce AP and makes it possible to breed heifers earlier and thus calve earlier.

5.6 Scrotal circumference EPDs and selection

Given its moderate to high heritability, the ease of measurement and the correlations with heifer AP and AFC, YSC is a trait of interest in beef cattle selection (Moser *et al.*, 1996). Not only the genetic correlation of YSC with AP and AFC is important, also its correlation with bull fertility traits. Several large American herdbooks, like the American Angus Association, the American Hereford Association and the North American Limousin Association, have already calculated EPDs for YSC (American Angus Association, 2016b; Hoagland, 2016). These EPDs are expressed in centimeters (cm) and predict the difference in YSC that will be passed to the bulls son. Bulls with a large YSC EPD are expected to give daughters who reach puberty earlier (Beef Improvement Federation, 2010; Breedplan, 2015).

Considering the moderate to high heritability estimates found for AP in heifers, it would be possible to select for AP directly, instead of using indirect selection by selecting for YSC (Moser *et al.*, 1996). But one of the main disadvantages of AP is that it is hard to measure (Martinez-Velázquez *et al.*, 2003). It requires, as mentioned above, several blood samples and laboratory analyses to identify the moment of puberty, which makes it in ranching conditions nearly impossible to measure. Another advantage of selection based on a sires YSC EPD, is that by using males it is possible to achieve a higher section intensity (Gargantini *et al.*, 2005). Given the positive genetic correlations of YSC with production traits (chapter 5.3), selection for a large YSC will not compromise production traits.

Right now (2017), breeders argue that the Black Hereford breed could benefit from decreasing the age at first calving (Hoagland, 2016). An EPD for YSC could help to achieve this goal. YSC EPDs have already proven their value in some breeds and studies. Moser *et al.* (1996) have set up an experiment to compare selection for YSC in a Limousin herd by phenotypic selection or by the use of EPDs. They found that selecting by using YSC EPDs is more effective than using phenotypic YSC measurements. By using YSC EPDs they were able to reduce the AP of heifers significantly.

The key of selection for a decrease in AFC and AP, is that there is an intermediate optimum. A late AP will result in heifers that calve late (> 2 years) and will lead to an economic loss. A premature AP will result in heifers that start cycling too young. They might come in heat before weaning, before they are separated from the bull calves. Moreover, it would break the yearly cycles that commercial cattlemen use in ranging conditions, where calving starts in late winter and ends in the spring (see appendix 1 for more information on the ranching conditions at J&N Ranch). Calves would get born in winter condition, when it is harder to feed the cows for a sufficient milk production (Hoagland, 2016).

One way of dealing with this intermediate optimum in selection is by implementing YSC in a selection index. A selection indexes is used to combine multiple traits into one single parameter to evaluate. Martin-Collado *et al.* (2016) studied the efficiency of different types of selection indexes for non-linear traits. They concluded that the use of a linear selection index, whose weighing factors are regularly updated, can be very efficient in reaching the optimal level of a trait. They also suggested another type of selection index which uses a non-linear quadratic function as weighing factor for its traits. This non-linear selection index has been found to have an equal efficiency as the linear updated index. Both types of selection indexes can be used for implementing YSC as a trait with an intermediate optimum in a breeding program.

6 Weight traits in beef cattle

6.1 Birth weight, weaning weight and yearling weight

Three important growth parameters in cattle breeding are the birth weight (BW), weaning weight (WW) and yearling weight (YW) measurements.

BW is defined as the weight measurement at birth (Beef Improvement Federation, 2010). The BW is important in calving ease (e.g. the relationship of BW with the weight of the dam) (Casas *et al.*, 2012). BW can be influenced by the sex of the calf, the season of birth and the age of the dam (Beef Improvement Federation, 2010).

WW is defined as the weight measurement at the moment of weaning. This measurement is often standardized to an age of 205 days, although it can be recorded between 150 and 250 days of age (Beef Improvement Federation, 2010). WWs are used to evaluate the differences in growth potential of the calves (Beef Improvement Federation, 2010) but it is also an indication for the milking ability of the dam. It can be influenced by the period of birth, the age of dam and the sex of the animal (Beef Improvement Federation, 2010). The average WW of beef cattle lies between 180 kg and 275 kg (Meyer (1997); Arthur *et al.* (2001); Torres-Vàquez *et al.* (2016)). Anderson (1977) found a significant effect of the age of dam on WW.

The YW is by definition the measured weight at 365 days of age. In commercial conditions this measurement has to be taken between 300 days and 550 days of age and has to be adjusted to an age of 365 days (Beef Improvement Federation, 2010). The YW is an important trait because of the genetic association with carcass weight and the efficiency of weight gain (Beef Improvement Federation, 2010; Torres-Vázquez & Spangler, 2016). YW can be influenced by the age of dam, the sex of the animal and the level of nutrition (e.g. grass-fed or a high energy diet) (Beef Improvement Federation, 2010). The average YW lies between 360 kg and 450 kg in beef cattle (Arthur *et al.* (2001); Torres-Vàquez *et al.* (2016)).

For these weight traits, it is important that the age at measurement is recorded. This makes an adjustment by age possible. If not, calves that are born late in the calving season can be given an unfair disadvantage over the older calves.

6.2 The heritability of weight traits

Most of the time, heritabilities for weight traits are found to be moderate to high (> 0.35). The following table (6.1) gives an overview of some of the heritability estimates found in the Black Hereford breed and in other beef breeds.

Table 6.1 A list of heritability estimates (h^2) for birth weight (BW), weaning weight (WW) and yearling weight (YW) in Angus, Hereford and Black Hereford cattle herds

Author	Year	Breed	h^2 BW	h^2 WW	h^2 YW
Smith *et al.* (a)	1989	-	0.27	0.14	0.29
Kriese *et al.*	1991	Hereford	0.45	0.33	0.50
Bennett and Gregory [a]	1996	Hereford	0.54	0.23	0.27
		Angus	0.26	0.25	0.42
Arthur *et al.* [a]	2001	Angus	-	0.17	0.28
Riley *et al.*[a]	2013	Black Hereford	0.28	0.35	0.23
Torres-Vàzquez [a]	2016	Hereford	-	0.35	0.36

[a] Some of the used animal models include a maternal genetic effect
Values indicated by a (-) were not reported or included in the study.

6.3 The genetic correlations between weight traits

The following table (6.2) gives an overview of some genetic correlations found in literature between BW, WW and YW.

Table 6.2 A list of genetic correlations found in literature between for birth weight (BW), weaning weight (WW) and yearling weight (YW) in Angus and Hereford cattle

Author	Year	Breed	BW - WW	BW - YW	WW - YW
Bourdon and Brinks	1986	Hereford	0.05	0.19	0.70
Smith *et al.* (a)	1989	-	0.25	0.41	0.84
Kriese *et al.*	1991	Hereford	0.32	-	-
Arthur *et al.*	2001	Angus	-	-	0.88
Torres-Vázquez	2016	Hereford	-	-	0.47

Values indicated by a (-) were not reported or included in the study.

Material and methods

7 Material and methods

7.1 Material

7.1.1 Observation records

In total, there are 2538 animals for which observations are available on one or more traits. Not every measurement is performed on every animal. The following table (table 7.1) summarizes measured traits and gives the number of measured animals for each trait. The measurements are performed at J&N Ranch (Leavenworth, Kansas, USA), except for 25 scrotal measurements, performed at another ranch. Although J&N ranch is the largest Black Hereford breeder, these records do not include the whole Black Hereford population.

Table 7.1 List of variables in the performance file of the Black Hereford cattle

Variable	Description	Sample size
Animal registration number	ABHA registration number of the animal	2538
Birth weight (BW)	Weight measured at birth (in kg)	2353
Wean Weight (WW)	Weight at the moment of weaning (in kg)	2079
Weaning date	Date when WW is measured	2070
Yearling weight (YW)	Weight at yearling age (in kg)	1448
Yearling date	Date when the YW and YSC are measured	1442
Yearling scrotal circumference (YSC)	Scrotal circumference at yearling age (in cm)	907

7.1.2 Pedigree records

The pedigree information is obtained from the American Black Hereford Association (ABHA) and the register from J&N Ranch. It contains 15152 records of animals born from 1980 till 2016. The records include the following information (table 7.2).

Table 7.2 List of variables in the pedigree file of the Black Hereford cattle

Variable	Description
Animal registration number	Registration number of the animal at the ABHA
Animal name	The official name of the animal, registered at the ABHA
Sire registration number	Registration number of the father of the animal at the ABHA
Dam registration number	Registration number of the mother of the animal at the ABHA
Birth date	Birth date of the animal
Sex	The sex of the animal

7.2 Methods

7.2.1 Data management and statistical analysis

The programming language R (version 3.3.2) and software environment Rstudio (version 1.0.136) are used as data management tool and to calculate several new variables. The general linear models are fitted by the *lm* function in R. The 'pedigree' package is used to manage and complete the pedigree file (Coster, 2013).

7.2.2 BLUP and REML estimation

The BLUPf90 family of programs is used for BLUP mixed model calculation and (co)variance estimation (Misztal *et al.*, 2015). This set of programs includes renumf90, blupf90, remlf90 and airemlf90 which are written in fortran90 (F90).

F90-programs are driven by a parameter file in which the model characteristics are specified. The renumf90 program reads the data- and pedigree file and creates from these files the appropriate input files for the other programs. Appendix 4 shows an example of a parameter file for renumf90. The airemlf90 program computes the BLUP of the animal model and meanwhile uses average information (AI)-REML to estimate the (co)variances. The first 5 computational rounds are computed by EM-REML and after these 5 rounds the program switches automatically to AI-REML, using the last estimate from EM-REML as a starting value for AI-REML. The AI-REML algorithm converges much faster than the EM-REML algorithm (e.g. 20 rounds in comparison to 1660).

Renumf90 selects a minimal pedigree from the whole pedigree file, constructed of the animals with records and their parentage. This pedigree consists of 4099 animals, of which 1567 are parents without records, and goes back as far as 6 generations. The solutions for the animal model include BLUP estimates for these 4099 animals.

7.2.3 The MCMC algorithm

The results of the animal model, fitted by BLUP, are compared to the results found by a MCMC algorithm. The used algorithm is the R function *MCMCglmm* from the R package 'MCMCglmm' (Hadfield, 2010). The animal model that is estimated, is the model that was found to be the most appropriate in the REML analysis. The calculation is only performed on the YSC trait, in order to save computation time.

7.2.4 Computer usage

The data editing, statistical analyses and AI-REML computations are performed on a personal computer. The more intense computations (the *MCMCglmm* function) are performed on the Linux high performance computer (Flemish Supercomputer Centre (VSC)).

7.2.5 Formulas used in validation and computation

7.2.5.1 Coefficient of determination

The coefficient of determination (R^2) is defined as the proportion of the residual variance in the model of interest, in comparison to the residual variance of the null-model (a model that only includes the intercept). This residual variance of the null-model is equal to the variance of the dependent variable (Y). The higher the R^2, the better the model explains the proportion of total variance. The R^2 is given by: (Xu, 2003; Nakagawa & Schielzeth, 2013).

$$R^2 = 1 - \frac{var(y_i - \hat{y}_i)}{var(y_i)} \tag{7.1}$$

with y_i the dependent variable of observation *i* and $\hat{y}_i$ the expected value for the dependent variable of observation *i*. The value $(y_i - \hat{y}_i)$ is called the residual value of the estimate.

7.2.5.2 Mean bias and mean squared bias

The mean bias (MB) of a model contains information about the accuracy or bias of the model. A MB larger than zero indicates a general overestimation of the values, a MB smaller than zero indicates an underestimation. The MB is given by the following formula (Kutner *et al.*, 2005; Tedeschi, 2006):

$$MB = \frac{Residuals}{n} = \frac{\sum(y_i - \hat{y_i})}{n} \tag{7.2}$$

where *n* is the number of observations, y_i is the i-th observation of trait y and $\hat{y_i}$ is the expected value for the i-the observation, given the model.

The mean squared bias (MSB) is an indicator for the precision of the estimates. A large MSB indicates a low precision and vice versa. The MSB is given by the following formula (Kutner *et al.*, 2005; Tedeschi, 2006):

$$MSB = \frac{Residuals^2}{n} = \frac{\sum(y_i - \hat{y_i})^2}{n} \tag{7.3}$$

7.2.5.3 *Genetic correlation*

The genetic correlation between two traits is calculated in the same way as the Pearson correlation and is based on the additive genetic (co)variances (Kutner et al., 2005):

$$r_{g\,(mn)} = \frac{\sigma^2_{a\,(mn)}}{\sqrt{\sigma^2_{a\,(m)}\,\sigma^2_{a\,(n)}}} \tag{7.4}$$

where $r_{g\,(mn)}$ is the genetic correlation coefficient between traits *m* and *n*, $\sigma^2_{a\,(mn)}$ the additive genetic covariance between traits *m* and *n* and $\sigma^2_{a\,(m)}$ and $\sigma^2_{a\,(n)}$ the respective additive genetic variances for traits *m* and *n*.

When selecting for a trait *m,* with a highly positive genetic correlation for trait *n*, trait *n* will also improve due to this selection. The reverse applies to traits with a negative genetic correlation.

7.2.5.4 *AIC*

The Akaike Information Criterion (AIC) is used in regression model choice (Akaike, 1974). It displays the amount of information that gets lost when a certain model is used. This way, it is a value for the accuracy of the model. It is possible to evaluate, based on the AIC value, multiple models and select the most suitable model. The most suitable model is the one with the lowest AIC value. The absolute value of the AIC has little interpretation. The AIC values is calculated by the following formula (Akaike, 1974):

$$AIC = 2k - 2\ln(L) \tag{7.5}$$

where k is the number of parameters to be estimated, and L is the maximized value of the likelihood function of the model.

7.2.5.5 The EBV, EPD and accuracy value

In the animal model, estimates are obtained for the effect of each individual in the dataset. These estimates of the additive genetic effect of the animals are also called the EBVs. The average EBV is zero and refers to the base population. EBVs are usually transformed into relative EBVs with the additive genetic value of the animals expressed relative to a chosen reference group of animals. A natural choice is to choose the active breeding animals as the reference group. In this analysis, the reference population consists of the 942 animals born from 2012 till 2016.

EPDs are calculated as the half of the obtained EBV.

As given in chapter 2.3.4, the accuracy value (r_i) of an EBV or EPD is equal to:

$$r_i = \sqrt{1 - \frac{PEV}{\sigma_a^2}} \tag{7.6}$$

where σ_a^2 is the additive genetic variance and the PEV (prediction error variance) is equal to the squared residual error of the EPD.

Results and Discussion

8 Results

8.1 Data management and descriptive statistics

The obtained data were checked for errors and corrected if possible. The variables sex and birth date from the pedigree file were added to the data file. The birth year of the animal was derived from the birth date of the animal. The age of the dam at the moment of calving was calculated from the birth date of the mother and the birth date of the calf. Dams older than 7 years at the moment of calving were merged with the group of 7-year old dams. The age at measurement of YW and YSC was calculated based on the birth day of the animal and the moment of the measurements. This trait was called yearling age (YA). The same was done for weaning age (WA). The variable 'ET' was added to the data for the animals born by embryo transfer. The YSC measurements were performed by three different observers. A separate variable 'observer' was added to the data. This resulted in a set of 2538 observations with the following variables (table 8.1):

Table 8.1 List of the variables available for analysis

Variable
Animal registration number
Yearling scrotal circumference (YSC)
Birth weight (BW)
Weaning weight (WW)
Weaning age (WA)
Yearling weight (YW)
Yearling age (YA)
Age of dam
ET
Sex
Birth year
Observer of YSC

Next, the data were screened for possible outliers. 1 outlier was found for BW (261 kg) and was deleted. 12 animals had a high YW (>550 kg) without a recorded YA. These animals were born in the same year (2001) and were probably weighted at an age older than 1 year. These 12 YW records were removed from the analysis. 2 animals had a YSC of 47 cm and 45 cm, respectively, without a record of the YA as covariate. These YSC measurement were removed. This resulted in observations on 2532 animals.

The following table (8.2) gives the number of measurements that are available for each level of the variables for the four traits. In total, there are 707 observations on YSC, 2239 on BW, 2053 on WW and 1411 on YW that are complete for every variable.

Table 8.2 List of the number of observations for every level of the variables for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW)

Trait	Level	YSC	BW	WW	YW
Age of dam					
(years)	2	161	540	466	293
	3	175	440	411	278
	4	144	337	313	228
	5	114	268	254	178
	6	97	216	204	133
	7 and older	195	438	412	311
ET					
	Yes	30	60	62	48
	No	864	2290	2017	1386
Sex					
	F	0	1222	1032	604
	M	894	1124	1043	830
Birth Year					
	1992	0	1	0	0
	1993	0	1	0	0
	1997	0	2	2	1
	1998	0	1	0	0
	1999	0	34	1	0
	2000	0	44	9	0
	2001	1	61	51	0
	2002	10	75	34	15
	2003	22	90	85	26
	2004	11	145	100	88
	2005	29	110	96	40
	2006	43	121	118	43
	2007	41	170	154	55
	2008	48	162	130	124
	2009	90	156	149	134
	2010	66	158	151	128
	2011	102	167	160	134
	2012	155	168	164	147
	2013	56	150	147	145
	2014	60	159	158	122
	2015	78	176	175	148
	2016	82	198	193	84
Observer					
	1	649	0	0	0
	2	220	0	0	0
	3	25	0	0	0

Table 8.3 gives the descriptive statistics of the four traits and table 8.4 gives the Pearson correlations between these traits (Kutner et al., 2005). Figure 8.1 visualizes the measurements in a scatterplot matrix.

Table 8.3 The descriptive statistics for yearling scrotal circumference (YSC) (in cm), birth weight (BW) (in kg), weaning weight (WW) (in kg), yearling weight (YW) (in kg), weaning age (WA) (in days) and yearling age (YA) (in days) measured in Black Hereford cattle from 1997 to 2016

	Sample size	Mean	Variance	Standard deviation	Min	Median	Max
YSC	894	33.2	7.9	2.8	26.5	36.3	44.6
BW	2350	37.5	16.4	4.0	22.8	38.2	51.0
WW	2077	245.6	1191.5	34.5	123.1	245.6	355.9
YW	1434	380.3	3578.2	59.8	205.8	384.3	588.4
WA	2068	201.3	473.4	21.8	106	201	283
YA	1426	336.3	858.7	29.3	249	331	435

Table 8.4 The Pearson correlations between yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW), weaning age (WA) and yearling age (YA) measured in Black Hereford cattle from 1997 to 2016

	BW	WW	YW	WA	YA
YSC	0.167	0.456	0.482	0.122	0.198
BW		0.253	0.261	-0.223	-0.054
WW			0.662	0.286	0.270
YW				0.166	0.548
WA					0.586

The average daily gain between the moment of weaning and yearling age is equal to 0.958 kg/day.

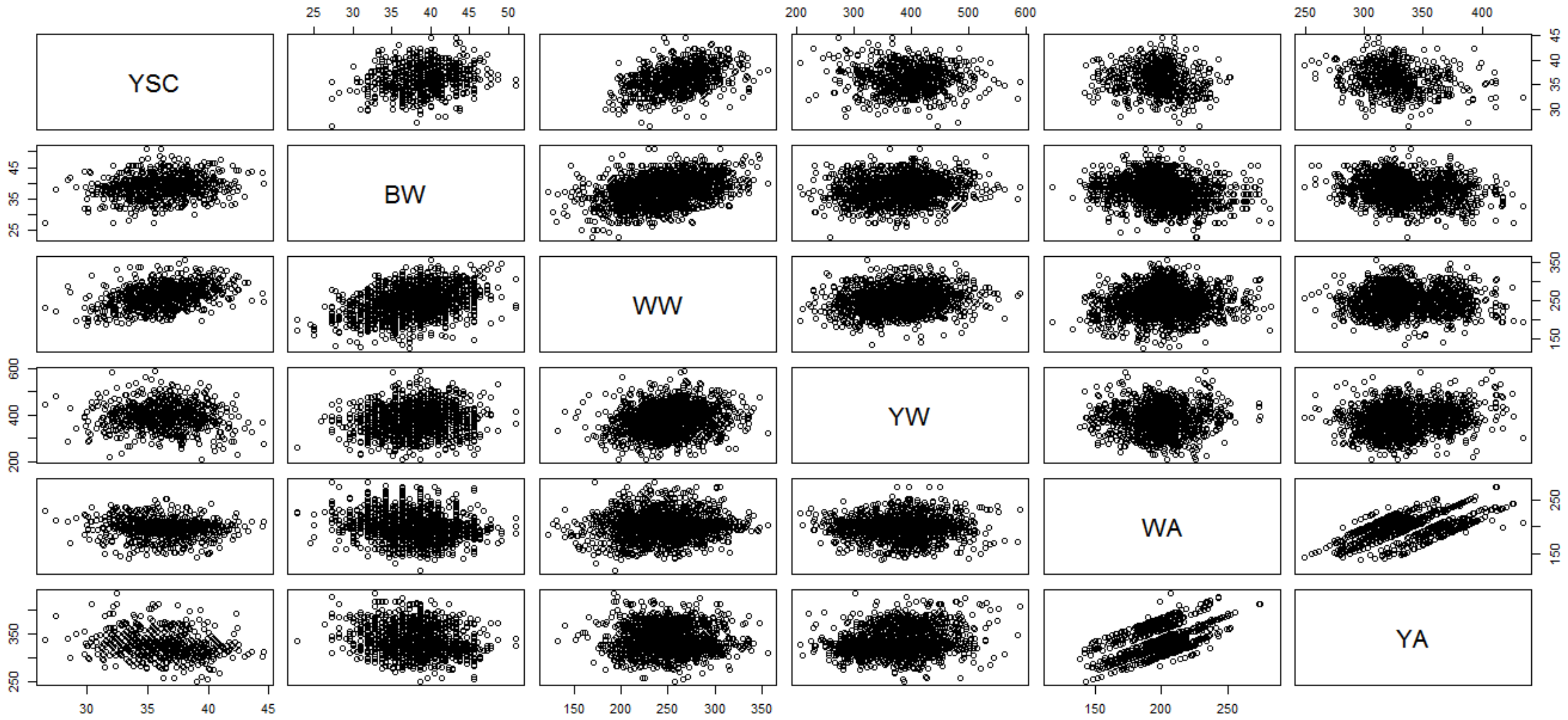

Figure 8.1 Scatterplot matrix for yearling scrotal circumference (YSC) in cm, birth weight (BW) in kg, weaning weight (WW) in kg and yearling weight (YW) in kg, weaning age (WA) in days and yearling age (YA) in days, measured in Black Hereford cattle from 1997 to 2016

8.2 The age at first calving

The age at first calving (AFC) is calculated from the pedigree data. This is done by first calculating the age of the dam for every calf. Then, the data of the calves from heifers are combined. This results in AFC data on 2683 animals. The mean AFC is equal to 955 days and the median 769 days, with a standard deviation of 336 days. Figure 8.2 shows the AFC frequency histogram for the dams with an AFC below 2000 days (5.5 years). Figure 8.3 gives more information about the AFC trait on each year of birth. For some cows, the AFC was calculated to be higher than 2000 days (excluded from figure 8.2 and 8.3).

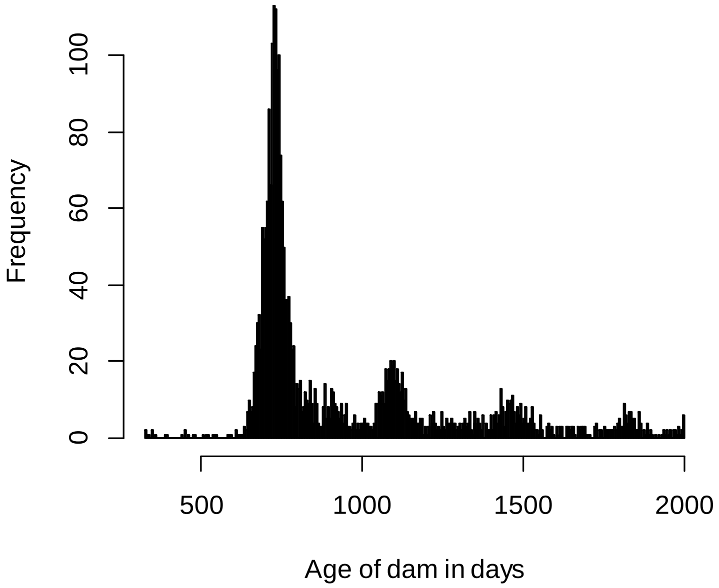

Figure 8.2 Frequency histogram of the age at first calving (in days) of cows in the Black Hereford registry from 1992 to 2016. Each bar has a width of 5 days.

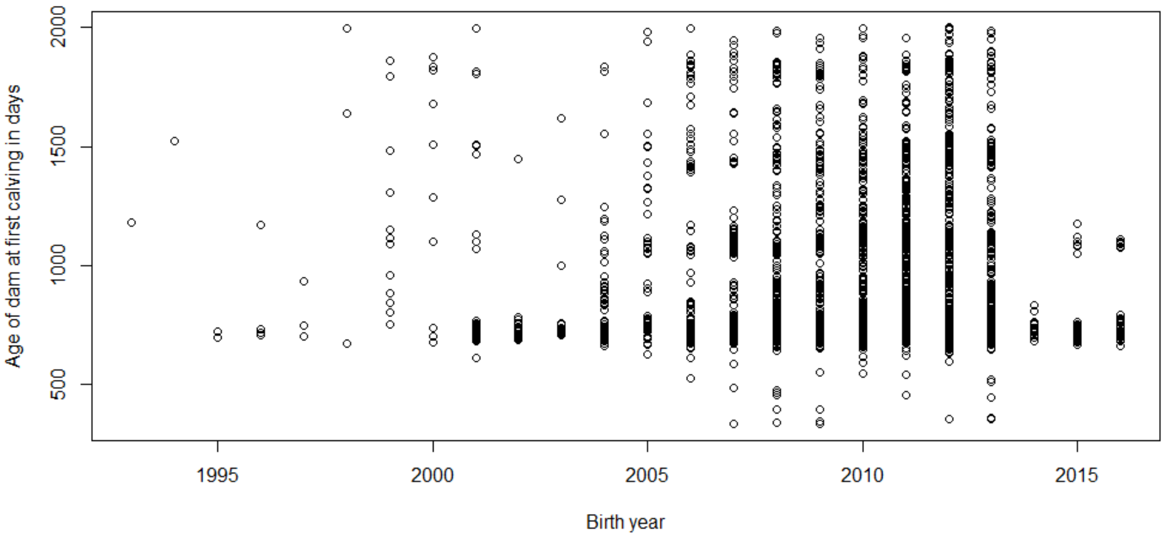

Figure 8.3 Plot of the age of dam at first calving (in days) by birth year in the Black Hereford registry from 1992 to 2016

8.3 Dams and sires versus birth year

One other thing that is evaluated, is whether sires and dams have progeny in more than one year. If not, the additive genetic effect of the sire or dam could be confounded with the birth year. If animals have progeny over more than one year, the entanglement of birth year effects and additive genetic effects decreases.

The 2532 recorded animals descend from 146 sires. On average, every sire has 17 descendants (with a standard deviation of 28) and has progeny in 2 different years (with a standard deviation of 1.4 years, ranging from 1 to 9 years). 91 sires only have descendants born in one single year.

From the 2532 recorded animals, there are 900 dams which averaged giving birth to 2.8 calves each (with a standard deviation of 2.5 calves). They are on average active in 2.8 years (with a standard deviation of 2.4 years, ranging from 1 to 12 years). 426 cows have only been active for one single year.

8.4 Linear regression results

Least squares regression is used to determine the significant independent variables for the four traits. In four univariate models, the YSC, BW, WW and YW traits are used as dependent variables and the variables found in table 8.1 are tested for their significance by a F-test (Kutner *et al.*, 2005). In these models the variable 'sire' is included to approximate the additive genetic effect. The sire-effect will be replaced by an additive genetic effect in the animal model. Table 8.5 gives a summary of the significant independent variables found in this analysis, given a significance level of 0.05. The complete table of results of the least squares regression can be found in appendix 5.

Table 8.5 Summary of the significant effects in univariate least squares regression models with yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) as dependent variable

Dependent variable			
YSC	BW	WW	YW
Sire	Sire	Sire	Sire
Birth year	Birth year	Birth year	Birth year
Age of dam	Age of dam	Age of dam	Age of dam
	Sex	Sex	Sex
	ET	ET	
Yearling age		Weaning age	Yearling age

8.5 The animal model results from airemlf90

8.5.1 Description of the models

Table 8.6 on the next page describes the animal models that are fitted by BLUP in this analysis.

Model 1 uses an adjustment formula for YSC, WW and YW. This adjustment formula corrects for the age at which the trait is measured. The formula for the dependent variable Y and animal *i* is:

$$Y_{adj,i} = Y_{obs,i} - f(Age_i - \overline{Age}) \tag{8.1}$$

with $Y_{adj,i}$ the adjusted value for animal *i*, $Y_{obs,i}$ the observed value for animal *i*, f the age adjustment factor, Age_i the age at measurement of the animal and $\overline{Age}$ the mean age at measurement for variable Y. The adjustment factors for YSC, WW and YW are equal to 0.0420, 0.908 and 0.934, respectively. These adjustment factors were obtained by least squares regression (chapter 8.4).

In model 3, 4 and 5 the age variables (YA and WA) are centered to a mean of zero and scaled to a standard deviation of one:

$$YA_{i\,scaled} = \frac{YA_i - 336\,days}{29\,days} \tag{8.2}$$

$$WA_{i\,scaled} = \frac{WA_i - 201\,days}{22\,days} \tag{8.3}$$

Every unit of scaled YA is equal to 29 days, and every unit of scaled WA is equal to 22 days. The birth weight measurements cannot be adjusted by age since the exact date of measurement, if different from the date of birth, is not given.

In model 4, birth year is treated as a random effect and a fixed observer effect is included. This way it is possible to estimate the effects of birth year as a part of all possible birth year effects. There will be a variance estimated for the birth year effects.

Model 5 includes the observer effect and birth year as a fixed effect.

Table 8.6 Overview of the composition of the five animal models used in this analysis

	Trait	Age at measurement	Additive genetic effect	Birth year		Age of dam	ET	Sex	Observer
			(random)	(fixed)	(random)	(fixed)	(fixed)	(fixed)	(fixed)
Model 1	YSC	Pre-adjusted to YA	x	x		x			
	BW	-	x	x		x	x	x	
	WW	Pre-adjusted to WA	x	x		x	x	x	
	YW	Pre-adjusted to YA	x	x		x		x	
Model 2	YSC	as covariate	x	x		x			
	BW	-	x	x		x	x	x	
	WW	as covariate	x	x		x	x	x	
	YW	as covariate	x	x		x		x	
Model 3	YSC	scaled, as covariate	x	x		x			
	BW	-	x	x		x	x	x	
	WW	scaled, as covariate	x	x		x	x	x	
	YW	scaled, as covariate	x	x		x		x	
Model 4	YSC	scaled, as covariate	x		x	x			x
	BW	-	x		x	x	x	x	
	WW	scaled, as covariate	x		x	x	x	x	
	YW	scaled, as covariate	x		x	x		x	
Model 5	YSC	scaled, as covariate	x	x		x			x
	BW	-	x	x		x	x	x	
	WW	scaled, as covariate	x	x		x	x	x	
	YW	scaled, as covariate	x	x		x		x	

YSC = yearling scrotal circumference, BW = birth weight, WW = weaning weight, YW = yearling weight, WA = weaning age, YA = yearling age, ET = embryo transfer

8.5.2 Model validation

The following table (8.7) gives for the five models the AIC value and the coefficients of determination (R^2) for the four traits in the five models.

Table 8.7 AIC values and the R^2 values for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) for the five models

	AIC	R^2 YSC	R^2 BW	R^2 WW	R^2 YW
Model 1	42643	0.687	0.668	0.777	0.309
Model 2	41870	0.684	0.666	0.752	0.788
Model 3	41380	0.739	0.666	0.773	0.839
Model 4	41744	0.729	0.668	0.769	0.837
Model 5	41373	0.732	0.666	0.773	0.840

Table (8.8) gives an overview of the mean bias (MB) and the mean squared bias (MSB) for the traits in the five models.

Table 8.8 Mean bias (MB) and mean squared bias (MSB) of the five models for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW)

	MB_{YSC}	MB_{BW}	MB_{WW}	MB_{YW}	MSB_{YSC}	MSB_{BW}	MSB_{WW}	MSB_{YW}
Model 1	0.101	0.0001	0.109	0.237	2.544	5.440	301.199	2587.483
Model 2	0.126	-0.0005	0.061	1.036	2.576	5.473	334.547	793.563
Model 3	0.118	0.0046	0.239	2.040	2.129	5.470	305.541	605.587
Model 4	0.120	0.0043	0.216	1.310	2.212	5.450	311.155	610.678
Model 5	0.119	0.0045	0.236	2.032	2.173	5.478	305.408	604.893

Next, the Spearman correlations of the EPDs for the five models are calculated (Kutner *et al.*, 2005). These correlations give more information whether the models rank the animals in the same way or not. Pairs of models with a high correlation indicate that the EPDs of the animals are ranked in the same way. The following table (8.9) contains the Spearman correlations for the models of the EPDs for the whole population and the active sire population (every sire born from 2012 till 2016). This population of active sires consists of 312 animals.

Table 8.9 The Spearman correlations of the expected progeny differences between the five models for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW). The upper diagonal matrix gives the correlations for the whole population and the lower diagonal matrix those for the active sire population. Values larger than 0.950 are highlighted in grey

YSC		Model 1	Model 2	Model 3	Model 4	Model 5
	Model 1		0.783	0.638	0.639	0.650
	Model 2	0.903		0.805	0.809	0.816
	Model 3	0.922	0.952		0.997	0.998
	Model 4	0.922	0.951	0.999		0.999
	Model 5	0.923	0.952	0.999	0.999	
BW		Model 1	Model 2	Model 3	Model 4	Model 5
	Model 1		0.972	0.940	0.922	0.941
	Model 2	0.981		0.980	0.965	0.981
	Model 3	0.966	0.993		0.993	0.999
	Model 4	0.962	0.991	0.999		0.993
	Model 5	0.966	0.993	0.999	0.999	
WW		Model 1	Model 2	Model 3	Model 4	Model 5
	Model 1		0.862	0.951	0.846	0.859
	Model 2	0.919		0.901	0.894	0.905
	Model 3	0.967	0.933		0.994	0.996
	Model 4	0.968	0.929	0.999		0.997
	Model 5	0.967	0.933	0.999	0.999	
YW		Model 1	Model 2	Model 3	Model 4	Model5
	Model 1		0.210	0.181	0.191	0.180
	Model 2	-0.063		0.893	0.881	0.898
	Model 3	-0.133	0.933		0.995	0.998
	Model 4	-0.125	0.928	0.999		0.996
	Model 5	-0.132	0.933	0.999	0.999	

The following table (8.10) gives the Pearson correlations (Kutner *et al.*, 2005) between the observed and the estimated observations, and the correlations between the estimated observations mutually for the five models.

Table 8.10 Pearson correlations between the observed and the estimated observations and the estimated observations mutually between yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) for the five models. Values larger than 0.950 are highlighted in grey

YSC		Model 1	Model 2	Model 3	Model 4	Model 5
	Observed YSC	0.751	0.861	0.896	0.895	0.893
	Model 1		0.885	0.826	0.822	0.826
	Model 2			0.975	0.973	0.975
	Model 3				0.999	0.999
	Model 4					0.999
BW		Model 1	Model 2	Model 3	Model 4	Model 5
	Observed BW	0.850	0.847	0.848	0.850	0.848
	Model 1		0.998	0.997	0.995	0.997
	Model 2			0.999	0.998	0.999
	Model 3				0.999	0.999
	Model 4					0.999
WW		Model 1	Model 2	Model 3	Model 4	Model 5
	Observed WW	0.708	0.898	0.894	0.894	0.895
	Model 1		0.782	0.725	0.722	0.725
	Model 2			0.988	0.988	0.988
	Model 3				0.999	0.999
	Model 4					0.999
YW		Model 1	Model 2	Model 3	Model 4	Model 5
	Observed YW	0.649	0.894	0.920	0.920	0.920
	Model 1		0.813	0.774	0.772	0.774
	Model 2			0.976	0.975	0.976
	Model 3				0.999	0.999
	Model 4					0.999

8.5.3 Estimates of variances, covariances, genetic correlations and heritabilities

The following table (8.11) contains the additive genetic and residual (co)variance estimates and their standard error, for the four dependent variables estimated by model 3. The (co)variance estimates of the four other models can be found in appendix 6.

Table 8.11 Variance and covariance estimates for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) estimated by model 3. The first diagonal matrix gives the additive genetic (co)variances, the second diagonal matrix gives the residual (co)variances. The value between parantheses gives the standard error of the estimate

	YSC	BW	WW	YW
YSC	3.41 (0.66)	0.99 (0.56)	20.10 (4.89)	30.42 (7.24)
BW		5.80 (0.77)	17.11 (4.74)	33.38 (7.26)
WW			405.12 (52.81)	501.65 (69.93)
YW				784.12 (118.41)
YSC	3.60 (0.48)	0.69 (0.41)	16.76 (3.48)	29.47 (5.15)
BW		7.86 (0.49)	14.73 (3.03)	15.49 (4.75)
WW			464.58 (32.98)	413.50 (44.22)
YW				941.67 (79.15)

The genetic and residual correlations, as estimated by model 3, are given in table 8.12. The heritabilities (equation 2.4, chapter 2.2) are given on the diagonal of this table. The correlations and heritabilities, obtained by the other models can be found in appendix 6.

Table 8.12 Genetic correlations (upper diagonal matrix), heritabilities (on the diagonal, highlighted in grey) and the residual variances (lower diagonal matrix) for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) estimated by model 3

	YSC	BW	WW	YW
YSC	0.49	0.22	0.54	0.59
BW	0.13	0.43	0.35	0.49
WW	0.41	0.24	0.47	0.89
YW	0.51	0.18	0.63	0.45

8.5.4 Parameter estimates

The BLUEs of the effects for YSC and the weight traits, obtained in model 3, can be summarized as following:

- Figure 8.4 shows the additive effect of the age of dam on the four traits. The visualized effects are relative to two-year old dams.
- Figure 8.5 shows the additive effect of birth year from 2002 to 2016 on the four traits. The visualized effects are relative to the year 2002.
- The estimate for scaled YA as a covariate for YSC is equal to 0.82 cm for every 29 days.
- For BW, there is a slight difference (0.75 kg) between ET calves and non-ET calves. The birth weight of male calves is 1 kg higher in comparison to female calves.
- The WW of ET calves is about 25 kg lower in comparison to non-ET calves. Male calves weigh on average 19 kg more than female calves at weaning. The scaled WA estimate as a covariate for WW is equal to 18.57 kg for every 22 days.
- Male animals weighted about 70 kg more than female animals at yearling age. The scaled YA covariate is estimated at 27.23 kg for every 29 days.

The airemlf90 program does not formally test whether these estimates are significantly different from zero.

In appendix 7, an overview is given for the estimated effects in the five models.

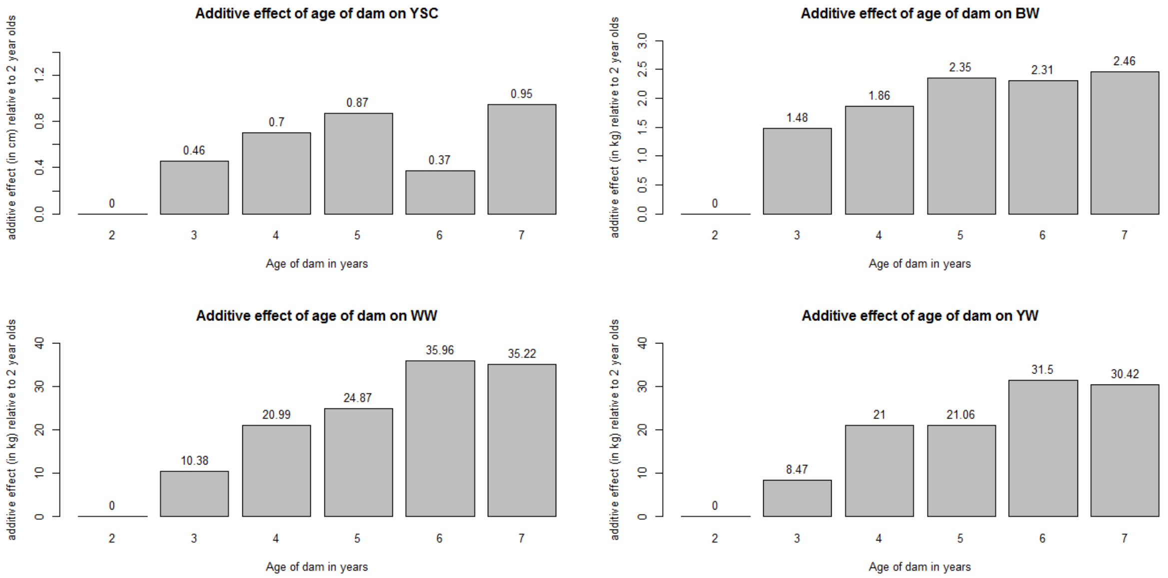

Figure 8.4 The additive effect of the age of dam on yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) as estimated by model 3

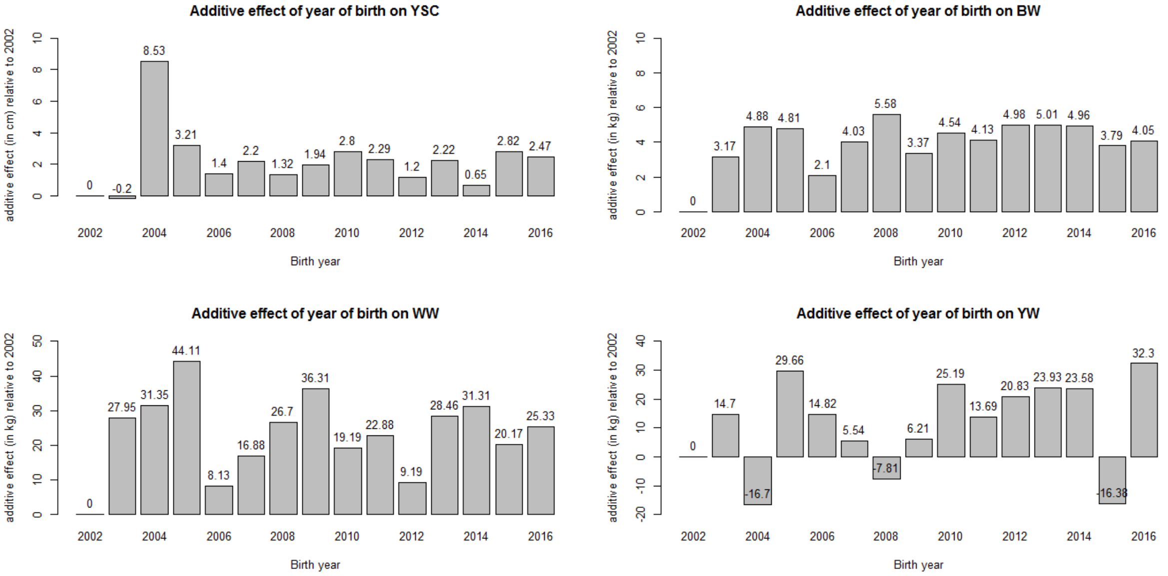

Figure 8.5 The additive effect of the year of birth on yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) as estimated by model 3

8.5.5 EPDs

The BLUP estimates of model 3 are used to calculate the (relative) EPDs of the 4099 animals and their corresponding accuracy values. Table 8.13 gives an excerpt of the list of 4099 animals as an example. The number of first-degree offspring is given in the column 'Offspring'. The animal registration numbers are changed to a coded number for privacy reasons. Table 8.14 gives an overview of the means, variances and standard deviations of the EPDs for the four traits. The Spearman correlations (Kutner et al., 2005) between the EPDs of the four traits are given in table 8.15. Figure 8.6 gives the frequency histogram of the EPDs for the four traits.

Table 8.13 An excerpt of the EPD list, giving the EPDs for yearling scrotal circumference (YSC) (in cm), birth weight (BW) (in kg), weaning weight (WW) (in kg) and yearling weight (YW) (in kg), the respective accuracy values and the number of offspring for eight Black Hereford animals, based on the estimates of model 3

Animal	EPD.YSC	ACC.YSC	EPD.BW	ACC.BW	EPD.WW	ACC.WW	EPD.YW	ACC.YW	Offspring
1	0.421	0.742	0.859	0.736	8.411	0.754	11.229	0.753	0
2	0.855	0.780	-1.169	0.837	13.835	0.850	17.109	0.830	12
3	0.182	0.825	-1.586	0.844	0.478	0.854	-3.641	0.845	11
4	-1.115	0.809	-0.497	0.844	-5.735	0.855	-5.774	0.841	13
5	0.249	0.844	0.284	0.868	9.656	0.877	14.344	0.866	22
6	0.385	0.774	0.133	0.802	1.774	0.809	7.421	0.802	5
7	-0.557	0.762	0.878	0.759	-2.370	0.775	-3.546	0.775	0
8	0.683	0.746	0.712	0.747	4.735	0.763	7.474	0.763	0

Table 8.14 Overview of the mean, variance, standard deviation and range of the EPDs for yearling scrotal circumference (YSC) (in cm), birth weight (BW) (in kg), weaning weight (WW) (in kg) and yearling weight (YW) (in kg), estimated by model 3

	Mean	Variance	Standard deviation	Min	Max
EPD YSC	-0.369	0.247	0.497	-1.731	2.069
EPD BW	-0.119	0.300	0.548	-2.146	2.513
EPD WW	-1.997	24.865	4.986	-26.403	20.196
EPD YW	-3.860	52.188	7.224	-39.993	28.046

Table 8.15 The Spearman correlations between the EPDs for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW), as estimated by model 3

	EPD BW	EPD WW	EPD YW
EPD YSC	0.337	0.605	0.677
EPD BW		0.455	0.554
EPD WW			0.924

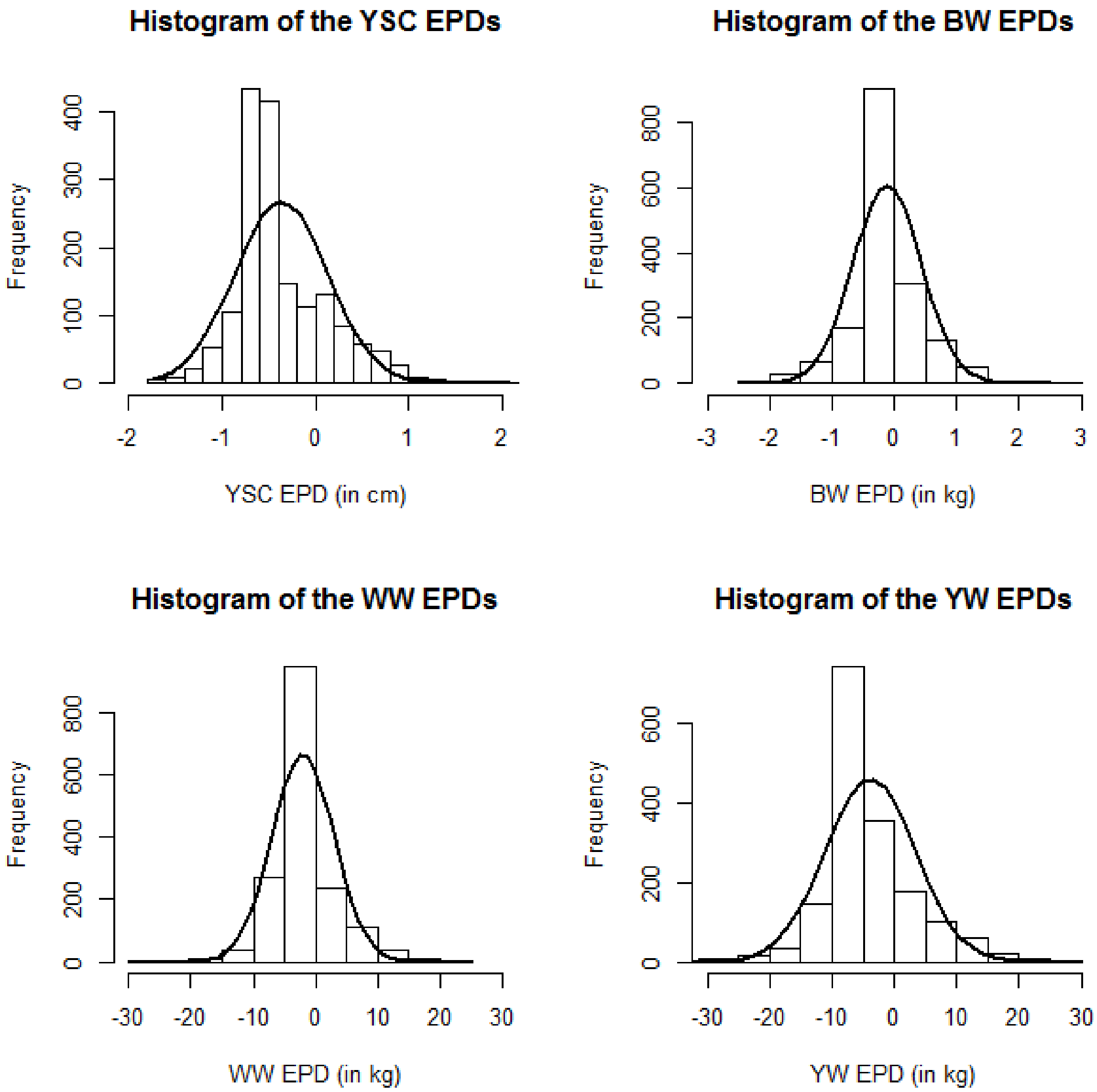

Figure 8.6 Frequency historgrams of the yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) EPDs estimated by BLUP. On each plot, a Normal distribution with the same mean and variance is plotted for comparison

8.6 The MCMC algorithm

A model with YSC as dependent variable, birth year, age of dam and scaled YA as fixed effects and the additive genetic effect as random effect (the analogue of model 3) is obtained by a MCMC algorithm. The 707 observations of YSC and the minimal pedigree of 4099 animals are used. The total computation time is 1 hour and 45 minutes. The Markov chain consists of 1 000 000 iterations, with a burn in period of 3000 and a thinning of 500 iterations. The *MCMCglmm* function uses for the fixed effects a wide Normal prior distribution (de Villemereuil, 2012). For the random and residual effect an inverse-Gamma distribution is chosen with *V* = 1 and *nu* = 0.002. The probability density function plot of this distribution can be found in chapter 2.3.6.5.

The behavior of the Markov chain is an important characteristic in the validation of the obtained results. This includes the behavior of the trace of the chain samples, the autocorrelation, the effective sizes and estimated posterior density function for each component of the model. The traces of the chain samples and the approximated posterior distributions for the estimated effects are given in appendix 8 (figure A8.1).

The model converges fast, which is typical for the *MCMCglmm* function (de Villemereuil, 2012). All approximated posterior distributions are symmetric and the traces show no clear trend. The autocorrelation between consecutive samplings of the additive genetic effects is 0.07. The effective sample size of the random additive genetic effects is 1708, for the residuals 1717. The effective sample sizes for the fixed effects can be found in table A8.1 in appendix 8.

The genetic variance of YSC is estimated at 2.21 cm², the residual variance at 3.66 cm² and the heritability of YSC amounts to 0.37.

The means of the approximated posterior distributions and the lower and upper values of the 95% highest posterior density intervals of the effects are given in appendix 8. The posterior means of the age of dam effects are given in figure 8.7. The 95% highest posterior density interval of the effect of the 6-year old dam includes zero. The highest posterior density intervals for the other levels of the age of dam do not include zero. So, the effect of age of dam is, except for 6-year old dams, significantly different from zero.

The posterior mean of the additive genetic effect of each animal is used to calculate a (relative) EPD for the animal.

The mean of the EPD estimates is equal to -0.044 cm, the variance 0.098 cm² and the standard deviation is 0.312 cm. Figure 8.8 gives the frequency histogram of the EPDs for YSC obtained with the MCMC algorithm.

The obtained EPDs are compared with the EPDs estimated by BLUP. The Spearman correlation between the two EPD estimations of the whole population is equal to 0.646 and the Pearson correlation is 0.725 (Kutner *et al.*, 2005). The Spearman correlation between the two EPD estimations for the 312 active bulls born from 2012 till 2016 is equal to 0.887 and the Pearson correlation is 0.877.

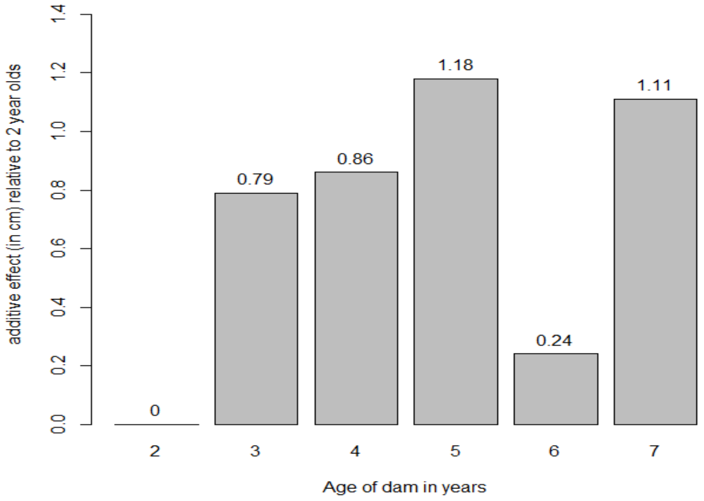

Figure 8.7 The mean of the posterior distribution of the age of dam on yearling scrotal circumference, relative to 2-year old dams, as estimated by the MCMC algorithm

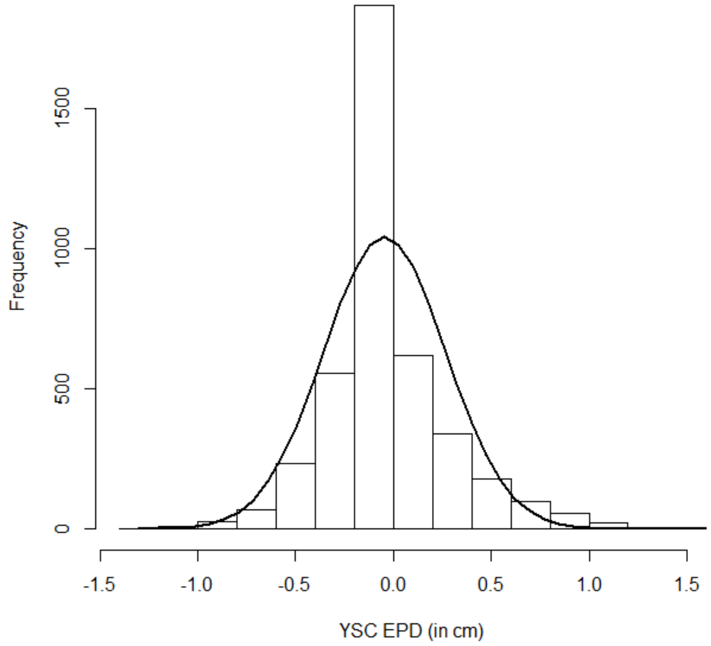

Figure 8.8 Frequency histogram of the yearling scrotal circumference, estimated by MCMC. A Normal distribution with the same mean and variance is plotted for comparison

9 Discussion

9.1 Data management and statistics

In total there are 894 observations for YSC. Only 707 of these are complete for every variable. 179 observations of the 894 do not have a YA record, 8 of the 894 do not include a record of the age of dam. When these measurements are missing, the observations cannot be taken into account by BLUP or the MCMC algorithm. Completeness of the pedigree file and data files is crucial for a good data analysis.

The average YSC of 35.99 cm is higher than most observations reported in literature (chapter 5.2). The standard deviation of 2.85 cm is conform the literature. The average BW, WW and YW is also in accordance with the averages found in literature (chapter 6.1). Note that the WA and YA is not recorded for every animal with a WW or YW measurement (9 and 18, respectively, missing).

The Pearson correlation found between YSC and WW, and YSC and YW is equal to 0.456 and 0.482, respectively. These are positive and indicate that the YSC increases for a higher WW or YW. Animals with a high YW will probably have a larger scrotum. This indicates that the growth of the scrotum of the animal is partly linked to overall growth of the animal. This association is previously found in other breeds (Bourdon & Brinks, 1986; Smith et al., 1989b). The correlation between YSC and WA, and YSC and YA is lower. The correlation between BW and WW, and BW and YW is low. The correlation between WW and YW is positive (0.662), which indicates that animals with a high WW will probably have a high YW and vice versa.

9.2 Age at first calving

The mean AFC is higher than expected in commercial conditions. In these conditions, it often occurs that heifers that do not calve before 3 years of age are culled. Many cows appear to have their first calf at an age of over 1100 days (± 3 years). The standard deviation (336 days) is high, indicating that the AFC is highly variable. This is also visible in the frequency histogram (figure 8.2).

The large AFCs are probably caused due to the way of calculating the AFC. There are no direct records of the AFC available, thus it has to be calculated as the difference between the birth dates of the cow and the first registered calf of the cow. One problem with this way of calculating is the fact that the first calf of a cow is not always registered. This is especially the case for the Angus and Hereford cattle that have been used in the initial crossings and for import of genetics. These animals are only registered at the moment they are used in the Black Hereford crossings, and there is no information about previous calvings transferred to the pedigree files of the Black Hereford breed. The last 3 years (2014 – 2016), the AFCs are more consistent (figure 8.3). During this period the Black Hereford breeding program temporarily stopped importing Hereford genetics into the breed (Hoagland, 2016).

In addition, AFC is highly influenced by management conditions. For example, J&N Ranch always synchronizes their heifers to be able to inseminate them at the same time (see appendix 1). This treatment blurs the genetic variance that might be present in the heifer population. Another management condition that influences the AFC is the moment of exposing the heifers to breeding.

Another thing that can be noted in figure 8.2 is that some animals have an age of dam below 600 days. These values are resulting from some inconsistencies in the birth dates of the animals.

For the above reasons, it was not worthwhile to include AFC in the further analysis.

9.3 Dam and sires vs birth year

Genetic links between the different birth years are observed coming from both paternal as maternal side. Sires appear on average in 2 years and dams in 2.8 years.

9.4 Linear regression results

The linear regression model shows the significant variables ($\alpha = 0.05$) that could be included in the animal model (table 8.5). The following independent variables are found to be significant:

- For YSC, the significant variables are: sire, birth year, age of dam and YA. The observer is not found significant, although this variable can be included in further analyses when more observers are involved.
- For BW, the significant variables are: sire, birth year, age of dam, sex and ET. For BW, there is no age adjustment possible, since the exact date at measurement, if different from the actual birth date, is not recorded.
- For WW, the significant variables are: sire, birth year, age of dam, sex, ET and the WA.
- For YW, the significant variables are: sire, birth year, age of dam and the YA. The ET effect has not been found significant.

9.5 Animal model results

9.5.1 Model description

Table 8.6 gives an overview of the used animal models. The results of the linear regression are used to decide which effects can be taken into account when constructing the models.

In model 1 YSC, WW and YW are pre-adjusted to the respective age at measurement. The adjustment factor for YSC is almost equal to the ones previously described in table 5.1. Those for WW and YW are close to the average daily gain from weaning to YA (chapter 8.1) and to the average daily gains reported for the Black Hereford breed in chapter 3.4. Model 2 estimates the animal model with unscaled ages at measurement. Model 3 estimates the same model, but with scaled ages. Model 4 estimates the birth year as a random effect. In order to do this properly, the observer effect is added to the model, although it is not found significant in the least squares regression. Without the observer effect, the birth year variance on YSC is estimated at 1471.5 cm². Adding the observer effect decreases the birth year variance of YSC to a more plausible 3.93 cm². This variance inflation is probably caused by the specific structure of the available data. Models 4 and 5 include the same effects but in model 5, the birth year effect is estimated as fixed.

In these five models the birth year effect can be interpreted as a contemporary group effect. It could be possible to use a birth year – period combination as a contemporary group, with period as a variable for the moment of birth in the year (e.g. the fall). This results in a large number of contemporary groups and many of these groups have only a few measurements. The accuracies of the estimates of the groups with only a few observations would be very low. Because of this, the birth year is chosen to be the contemporary group variable in the models.

9.5.2 Validation and selection of the different models

In table 8.7, not all AIC values can be compared with each other. Models 1 (with pre-adjusted measurements) and 2 (with unscaled ages) are estimated on different datasets than models 3, 4 and 5 (with scaled ages). Models 3, 4 and 5 are different mixed models, fitted on the same data. Their AIC values can be compared. Of these three models, the AIC values of 3 and 5 are lower than the value of model 4. The AIC value of model 5 is a little lower than the one from model 3. This is probably caused by the extra independent variable that is present in model 5. Because this model has an extra factor, it loses less information and therefore its AIC value is lower.

The R^2 values for the BW and WW traits are in the five models fairly alike (table 8.7). The main differences in R^2 between the five models can be found in the YSC and YW traits. For YSC, models 3, 4 and 5 explain more of the total variance in comparison to models 1 and 2. For YW, model 1 only explains 30.9% of the total variance, while model 2 explains 78.8% and models 3, 4 and 5 explain ± 84%.

The MBs in the five models for YSC are about equal, but all larger than zero. This indicates that, in general, the YSC trait is overestimated. The MB for BW is almost equal to zero in the five models. For WW, the MB is the highest (± 0.220) in models 3, 4 and 5 and is the lowest in model 2. For YW, the MB is the highest (± 2) in models 3 and 5, being equal to ± 1 in models 2 and 4 and equal to 0.237 in model 1.

The MSB for YSC is the lowest in models 3 and 5. A low MSB indicates that the estimates are more precise. The MSB for BW is equal in the five models. For WW, the MSB is the lowest in models 1, 3 and 5. For YW, the lowest MSBs are found in models 3, 4 and 5 (± 605) and model 1 has a MSB equal to 2587. This indicates that model 1 has a low precision for YW.

In general, the Spearman correlations of the EPDs (table 8.9) for the whole population are lower than these for the active sire population. The ranking of the EPDs of the animals in the active sire population is very similar in the five models. This indicates that the models value the active sire population more in the same way as they value the whole population. Model 1 stands out since it has lower correlations with the other models for YSC, WW and especially for YW. It values the active population for YW completely different in comparison to the other models. Model 3, 4 and 5 have for the four traits correlations almost equal to one, both for the whole as the active sire population. The ranking of the EPDs of the animals is almost equal, indicating that these three models value the breeding values of the animals in the same way.

The Pearson correlations (table 8.10) between the observed traits and the estimated observations give an indication which model is better capable to estimate the observed values. For YSC, WW and YW, the models 3, 4 and 5 do this equally good, with correlations of about 90%. Apart from this, table 8.10 gives also information about the Pearson correlations between the estimated observations for the five models mutually. Model 2, 3, 4 and 5 have almost equal estimated observations. Model 1 has lower correlations for YSC, WW and YW, indicating that the estimated values are different from the other models.

The validation parameters indicate that models 3, 4 and 5 estimate the EPDs in the same way. They have equal estimated observations, R^2 values that are almost equal and have MBs and MSBs that only vary a little. Model 1 estimates the EPDs and observations often different from the other models and has, especially for YW, a low R² and a high MSB. The R² values for model 2 are for YSC and YW lower than the ones estimated for models 3,4 and 5. Model 2 has also a higher MSB for YW. The five models estimate the BW trait almost equally.

Model 3 is chosen as the most appropriate model for the available data. Model 4 and 5 have a lower AIC value but they include the non-significant observer effect. The parameter and EPD estimates are almost equal for these three model, so either of these three models obtain the same results.

9.5.3 Estimates of variances, covariances, genetic correlations and heritabilities

The additive genetic variance and covariance estimates for model 3, 4 and 5 are almost equal. Same counts for their residual (co)variance estimates. Model 1 obtains lower estimates for both the additive genetic and residual (co)variances. The additive genetic variance for YW is ten times smaller than in other models. The estimates of model 2 are in between these of model 1 and model 3, 4 and 5.

The heritability estimates for model 3 are for the four traits moderate to high (> 0.35). The estimate for YSC is equal to 0.49 and is conform the estimates found in literature (chapter 5.2). The heritabilities of the weight traits in the Black Hereford breed are previously estimated by Dr. Riley *et al.* (0.28, 0.35 and 0.23 for BW, WW and YW, respectively). The heritabilities obtained in this analysis are somewhat higher (0.43, 0.47 and 0.45 for BW, WW and YW, respectively). This can be explained by the fact that the models in this analysis included other factors (e.g. another contemporary grouping structure). Moreover, Riley *et al.* included data from multiple herds and estimated a maternal genetic effect. When only one herd is measured, the managemental and environmental variability is limited. The observed differences are more likely to be caused by genetic differences. Including a maternal genetic effect decreases the additive genetic variance and thus decreases the (additive) heritability.

Both the heritability estimates for BW (0.43) and YW (0.45) are in accordance with values found in Angus and Hereford breeds as reported in chapter 6.2. The estimate for WW (0.47) is somewhat higher than those found in other Hereford and Angus herds (0.14 to 0.35, chapter 6.2).

The heritability estimate for YW in model 1 is clearly off. This is caused by a low additive genetic variance estimated when using the pre-adjusted YW measurements.

The heritability estimates for model 3 and 5 are almost equal, but those of model 4 differ. This is caused by the design of model 4 with the birth year effect as random effect. This creates an extra variance term in the model and thus in the heritability calculation, making the denominator larger and the heritability estimate lower.

The genetic correlations between YSC and BW, WW and YW are positive and, respectively, low (0.22), and moderate (0.54 and 0.59, respectively). These positive correlations are in accordance with the literature given in chapter 5.3. It indicates that the genetic relationship between YSC and growth traits found in other breeds, is also present in the Black Hereford breed. So, when selecting for a higher YSC, the WW and YW will also increase. The high genetic correlation (0.89) between WW and YW is conform the literature previously reported in chapter 6.3. Thus, when a breeder selects for a higher WW, the YW will also be positively influenced.

9.5.4 Parameter estimates

Models 2, 3, 4 and 5 have about the same estimates for the effects for the four traits. The estimates of model 1, except for BW, differ more from the other four models. The estimates for BW are in the five models almost equal.

9.5.4.1 Age of dam

As described in chapter 5.1, previous studies indicate that the age of dam is of influence for the YSC of their sons. The older the dam, the larger the bulls scrotum at yearling age. For the Black Hereford breed, the same trend is found. The linear regression in chapter 8.4 indicates that, for YSC, the age of dam is significant and the BLUEs in figure 8.4 show the increase in YSC in sons of older dams. For example, a son of a cow of 5 years old has on average a 0.87 cm larger YSC in comparison to a son from a heifer. The estimate for 6 year old dams is lower than those for 5- and 7 year old dams. There is no clear explanation for this decrease.

The influence of the age of dam on the YSC can be caused by the correlation between YSC and weight traits. Older cows, which are selected every year again in the breeding program, most probably have favorable growth characteristics. It is possible that the observed increase of YSC in older dams is caused by the selection of the dams for their superior weight traits. These cows will have offspring with higher WW and YW. Their sons can have a higher YSC, given the genetic correlation between YSC and WW, and YSC and YW.

The BLUEs for the age of dam effects on BW, WW and YW increase for older dams (figure 8.5). This can partly be attributed to the previously discussed selection effect or by the fact that the mothering ability of a cow increases in older cows (Smith *et al.*, 1989a).

9.5.4.2 Birth year

For YSC, the mean effect of birth year, excluding 2004, is equal to 1.7 cm. For 2004 there is a large effect estimated (8.53 cm). This is probably caused by the fact that in 2005 (for animals born in 2004) some animals were recorded with a much larger YSC (e.g. animals older than one year but with an incorrect YA). Indeed, of the 11 measurements of YSC on animals born in 2004, 8 have a YSC larger than 40 cm. Their YAs are around average. It is possible that these records are actually taken on older animals, without recording their real age at measurement.

For the weight traits, especially for WW and YW, the birth year effect is more variable. This can be caused by differences in annual forage availability in the pastures or by weather conditions. Note that model 4 estimates for every year an effect on WW and YW, although some of these years (e.g. 1992 and 1993) have no measurements for WW or YW. This is possible since birth year is estimated as a random effect.

9.5.4.3 Age at measurement

The age at measurement estimates are for the four traits, in the five models, larger than zero. This indicates that older animals will have a higher YSC, WW or YW. The age adjustment factor for YSC in model 3 is equal to 0.82 cm for every 29 days, given the used scaling in chapter 8.5.1. This results in an increase of 0.0283 cm for every day the bull is older than 336 days at the moment of measurement. This is more or less conform the values found in literature for other beef breeds (see chapter 5.1) and approximates the adjustment factor used in model 1 (0.0420 cm/day).

9.5.4.4 Observer

The effects of the observer for YSC in models 4 and 5 show only small differences between the three observers. This is expected, since the observer effect is not found to be significant (chapter 8.4).

9.5.4.5 ET

Animals born by ET have a lower BW and WW estimate in comparison to non-ET calves. This is somewhat unexpected because the used donor cows are expected to be the better cows in the herd. It is possible that the recipient cows provide not these optimal rearing conditions (in utero and during rearing).

9.5.4.6 Sex

For BW, WW and YW, female animals have a lower weight in comparison to the male animals.

9.5.5 EPDs

In this analysis, EPDs are estimated for 4099 animals. The means of the estimated EPDs are somewhat below zero (table 8.12). This is also visible in figure 8.6, where the histograms show for YSC and YW right - skewed distributions. It is clear that a lot of the EPDs are estimated somewhat below zero. One reason for the high number of EPDs below zero could be the large number of animals with only a limited amount of information. These animals have an EPD close to zero and a low accuracy. Because of the transformation of the EPDs to the reference group (as given in chapter 7.2.5.5), this group of animals might have received a negative EPD.

Table 8.13 gives the calculated EPDs and accuracies of eight animals as an excerpt of the whole EPD list. For example, animal 1 has an EPD for YSC of 0.421 cm. This indicates that the male progeny of this animal, if mated to an animal with an EPD for YSC of 0 cm, is expected to have an YSC of 0.421 cm larger than the population average. A calf of animals 1 and 2 from table 8.13, if both animals are of different sexes, is estimated to have a birth weight 0.31 kg lower than the population average, its WW is expected to be 22.2 kg higher than the population average and its YW 28.3 kg higher.

The accuracies of these eight animals in table 8.13 are fairly high. For these eight animals, it is clear that animals with offspring have an higher accuracy in comparison to animals without offspring. When evaluating the whole table of EPDs and accuracies (not given), most animals with a high accuracy value are either animals with a large number of offspring or animals with observations of their own. Animals with only a few relatives and no observations of their own have a low accuracy (0.10 – 0.30).

The Spearman correlations between YSC and WW, and YSC and YW are moderate (0.605 and 0.677, respectively) (table 8.15). This indicates that the ranking of the animals by YSC EPDs is for 60.5 % equal to the ranking of the WW EPDs, and 67.7 % equal to the ranking of the YW EPDs. When selecting for an animal with a high YSC EPD, it is likely that this animal also has a high WW and YW EPD. The Spearman correlations between WW and YW is almost equal to one (0.924). This implies that the EPDs of the animals for both traits are ranked in the same order. Animals with a high WW EPD will probably also have a high YW EPD and vice versa.

9.5.6 A remark on model 1

It seems that pre-adjusting the dependent variables for their age at measurement, as done in model 1, is not optimal. This adjustment is highly dependent on the used adjustment factor. Model 1 uses linear adjustment factors, obtained by least squares regression. For YSC, the estimated effects and (co)variances are quite equal to the ones for the other models. But in YW there is a large difference in the obtained estimates in comparison to the other models.

Other linear adjustment factors have been tried but did not improve the results (results not given). These other adjustment factors were obtained from the covariate estimates in model 2: 0.0043, 0.46 and 0.21 for YSC, WW and YW, respectively.

It is possible that a linear adjustment factor is not sufficient. An alternative adjustment factor could be a calculation based on the average growth of the animal between the moment of weaning and yearling age. This makes the adjustment more robust against individual variation in growth rates.

9.6 The MCMC algorithm

The MCMC computation time of the animal model is higher than the computation time of the airemlf90 program. The BLUP computations finishes in minutes, where the MCMC algorithm needs over an hour to get results. For this reason, only a single trait model is estimated, instead of the four-trait model fitted by BLUP. The main drawbacks of the MCMC algorithm in comparison to the BLUP computations are the high computation time and the fact that MCMC is less user-friendly (e.g. setting the prior distributions, checking the convergence and high autocorrelation). Same drawbacks are found by de Villemereuil *et al.*, (2013).

For the MCMC calculation, the biggest problem is the large autocorrelation between consequent runs. This large autocorrelation is caused by the low number of observations (707) in comparison to the number of animals to estimate (4099). This results in a large thinning interval. In order to obtain an acceptable effective sample size, the total number of iterations is increased. This large number of iteration lengthens the computation time.

The behavior of the algorithm is checked before looking at the results of the computation (see appendix 8 for the trace plots and the approximated posterior distributions of the effects). The traces of the chains do not show a trend and the estimated posterior density functions look symmetric. Convergence has occurred and the chains do not get “stuck” on a specific value.

An autocorrelation of 0.07 is acceptable (de Villemereuil, 2012). It points out that the correlation between consecutive samples is estimated to be 7 percent. The obtained effective sample size for the additive genetic effect (1708) is moderate. Sizes above 1000 are acceptable (de Villemereuil, 2012). The effective sample sizes of the fixed effects are larger than those of the residuals and the additive genetic effect. A high autocorrelation reduces the effective sample size and a high effective sample size is necessary to obtain a high number of independent variables to estimate the posterior distributions.

The genetic variance of YSC is estimated at 2.21 cm^2. This is lower than the estimate from model 3 in the BLUP computation (3.41 cm^2). The heritability estimate (0.37) is lower than the ones estimated by BLUP but is still in accordance with the estimates previously reported in chapter 5.2.

The Spearman correlations between the EPDs estimated by BLUP and those obtained by the MCMC algorithm are moderate for the whole population (0.646) and high for the active sire population (0.887). This indicates that both computations result for the active sire population in roughly the same ranking of the animal. The Pearson correlation between the EPDs obtained by BLUP and by MCMC indicate that the EPD values in both computations are fairly alike (0.725 and 0.877, respectively for the whole population and the active sire population).

The effect of the age of dam on YSC (figure 8.7) is the same as the one observed in the BLUP computation. The 95% highest posterior density intervals of the effects of age of dam, except for 6-year old dams, do not include zero. For sons of 6-year old dams, the same decrease in YSC can be observed as found in the BLUP estimation. The estimated birth year effects are comparable to the ones found in the BLUP computations. The posterior mean of the YA effect is estimated at 1.17 cm for 29 days. This value is equal to +0.0403 cm per day and is in accordance with the values previously described in chapter 5.

Conclusion

10 Conclusion

The preferred model in this analysis is model 3 and includes:

- For YSC: birth year and age of dam as fixed effect, the scaled YA as covariate and an additive genetic effect as random effect.
- For BW: birth year, age of dam, ET and sex as fixed effects and an additive genetic as random effect.
- For WW: birth year, age of dam, ET and sex as fixed effect, the scaled WA as covariate and an additive genetic effect as random effect.
- For YW: birth year, age of dam and sex as fixed effect, the scaled YA as covariate and an additive genetic effect as random effect.

Birth year can also be included as a random effect (see model 4). When the number of observers increases in future analyses, an observer effect can be added to the animal model (see model 5).

This four-trait model is fitted by BLUP with REML (co)variance estimation. Also, a MCMC algorithm is used to evaluate the YSC separately. Both methods can be used in animal breeding but MCMC has some drawbacks over BLUP (e.g. a high computation time for complex multi-trait models).

The heritabilities of YSC, BW, WW and YW in Black Hereford cattle are moderate to high (0.49, 0.43, 0.47 and 0.45, respectively). This indicates that selection for YSC and weight traits is possible. This is the first time that the heritability for YSC is estimated in the Black Hereford breed.

EPDs for YSC, BW, WW and YW weight are estimated for 4099 Black Herefords. EPDs for weight traits were previously reported for the Black Hereford breed but this is the first time EPDs are estimated for YSC.

Due to management conditions and data inconsistencies, a possible link between YSC and AFC in the Black Hereford breed could not be studied. In order to be able to study this, certain criteria on keeping records for the AFC of animals will have to be imposed in agreement with the breeders. To measure the AFC directly, breeders have to change their management conditions (e.g. no longer estrus synchronizing the heifers and exposing heifers to natural breeding). In that case, the birth date of the calf is a representation of the AFC of the mother.

To improve the accuracy of the heritability, genetic correlations and the EPD calculations, it is important that measurements are continued. The measurements can be extended to other herds. When these measurements are performed in other herds, it is important that all possible variables (e.g. breeding groups, observers, management conditions) are recorded.

References

11 References

Acquaah, G. (2012). *Principles of Plant Genetics and Breeding*. *Journal of Chemical Information and Modeling* (Second edi, Vol. 53). Bowie State University, Maryland, USA: Wiley-Blackwell.

Akaike, H. (1974). A new look at the statistical model identification. *Transactions on Automatic Control*, *19*(6), 716–723.

American Angus Association. (2016a). *2016 Annual Report*. Saint Joseph MO, USA: American Angus Association.

American Angus Association. (2016b). Personal Interview. Saint Joseph MO, USA: (personal communication).

American Angus Association. (2016c). *Sire evaluation report, fall 2016*. Saint Joseph MO, USA.

American Hereford Association. (2016). *2016 American Hereford Association Annual Report*. Kansas City MO, USA: American Hereford Association.

Anderson, J. H. (1977). *Factors affecting weaning weights of beef cattle*. Iowa State University.

Arthur, P. F., Archer, J. A., Johnston, D. J., Herd, R. M., Richardson, E. C., & Parnell, P. F. (2001). Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *Journal of Animal Science*, *79*, 2805–2811.

Beef Improvement Federation. (2010). *Guidelines For Uniform Beef Improvement Programs*. Raleigh NC, USA: Joe Cassady, North Carolina State University.

Berry, D. P., & Evans, R. D. (2014). Genetics of reproductive performance in seasonal calving beef cows and its association with performance traits. *Journal of Animal Science*, *92*, 1412–1422.

Berry, D. P., Wall, E., & Pryce, J. E. (2014). Genetics and genomics of reproductive performance in dairy and beef cattle. *Journal of Animal Science*, *8*(1), 105–121.

Bourdon, R. M., & Brinks, J. S. (1986). Scrotal circumference in yearling Hereford bulls: adjustment factors, heritabilities and genetic, environmental and phenotypic relationships with growth traits. *Journal of Animal Science*, *62*, 958–967.

Bouron, W. F., & Boulpaep, E. L. (2003). *Medical Physiology* (1st ed.). Philadelphia PA, USA: Saunders.

Breedplan. (2015). *A basic guide to breedplan EBVs*. Armidale NSW, Australia: Breedplan C/- ABRI University of New England.

Brinks, J. S., McInerney, M. J., & Chenoweth, P. J. (1978). Relationship of age at puberty in heifers to reproductive traits in young bulls. In *Proc. West. Sect. Am. SOC. Anim. Sci. 29:28*.

Buchanan, D., & Hanna, L. (2014). Understanding EPDs. *Black Hereford Journal*, *3*(1), 7–10.

Bullock, D. (2009). Fundamentals of Expected Progeny Differences, (June), 1–4.

Burns, B. M., Gazzola, C., Holroyd, R. G., & Crisp, J. (2011). Review Article Male Reproductive Traits and Their Relationship to Reproductive Traits in Their Female Progeny : A Systematic Review. *Reproduction in Domestic Animals*, *46*, 534–553.

Cammack, K. M., Thomas, M. G., & Enns, R. M. (2009). REVIEW : Reproductive Traits and Their Heritabilities in Beef Cattle. *The Professional Animal Scientist*, *25*, 517–528.

Casas, E., Thallman, R. M., Cundiff, L. V, Casas, E., Thallman, R. M., & Cundiff, L. V. (2012). Birth and weaning traits in crossbred cattle from Hereford, Angus, Brahman, Boran, Tuli and Belgian Blue sires. *Journal of Animal Science*, *89*, 979–987.

Coster, A. (2013). Pedigree: Pedigree functions. Retrieved from https://cran.r-project.org/package=pedigree.

Day, M. L., & Nogueira, G. P. (2013). Management of age at puberty in beef heifers to optimize efficiency of beef production. *Animal Frontiers*, *3*(4), 6–11.

de Villemereuil, P. (2012). Tutorial Estimation of a biological trait heritability using the animal model How to use the MCMCglmm R package.

de Villemereuil, P., Gimenez, O., & Doligez, B. (2013). Comparing parent – offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits, 260–275.

Denton, A. S. (2009a). Believing in Heterosis. *The Whiteface: Journal of the American Hereford Association*.

Denton, A. S. (2009b). Building the baldie. *The Whiteface: Journal of the American Hereford Association*.

Ekarius, C. (2008). Storey's illustrated breed guide for sheep goats cattle and pigs. Storey.

Evans, J. L., Golden, B. L., Bourdon, R. M., & Long, K. L. (1999). Additive Genetic Relationships Between Heifer Pregnancy and Scrotal Circumference in Hereford Cattle. *Journal of Animal Science*, *77*, 2621–2628.

Felius, M. (1995). *Cattle breeds an encyclopedia*. Doetinchem, NL: Misset Uitgeverij.

Felius, M. (2016). *On the breeds of cattle: Their history, classification and conservation*. University of Utrecht.

Franke, D. E., Burns, W. C., & Koger, M. (1975). Variation in Coat-Color Pattern of Hereford cattle. *Journal of Heredity*, *66*(3), 147–150.

Gargantini, G., Cundiff, L. V, Lunstra, D. D., & Van Vleck, L. D. (2005). Genetic Relationships Between Male and Female Reproductive Traits in Beef Cattle Genetic Relationships Between Male and Female Reproductive Traits in Beef Cattle. *The Professional Animal Scientist*, *21*, 195–199.

Gilmour, A. R., Thompson, R., & Cullis, B. R. (1995). Average Information REML: An efficient algorithm for variance parameter estimation in Linear mixed models. *Biometrics*, *51*, 1440–1450.

Greiner, S. P. (2009). Understanding Expected Progeny Differences (EPDs).

Grosz, & MacNeil. (1999). The "spotted" locus maps to bovine chromosome 6 in a Hereford-Cross population. *Journal of Heredity*, *90*(1), 233–236.

Gugelmeyer, S. (2009). Pennsylvania producers add value with feeder calf pool. *The Whiteface: Journal of the American Hereford Association*.

Gutiérrez, J. P., Álvarez, I., Fernández, I., Royo, L. J., Díez, J., & Goyache, F. (2002). Genetic relationships between calving date, calving interval, age at first calving and type traits in beef cattle. *Livestock Production Science*, *78*, 215–222.

Hadfield, J. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*, *33*(2), 1–22.

Hadfield, J. (2016). *MCMCglmm Course Notes*.

Hahn, J., Foor, H., & Seidel, G. E. (1969). Testicular growth and related sperm output in dairy bulls. *Journal of Animal Science*, *29*, 41–47.

Henderson, C. R. (1984). Prediction of Random Variables. In L. R. Schaeffer (Ed.), *Applications of Linear Models in Animal Breeding* (3rd ed.). Guelph, ON: University of Guelph.

Hoagland, J. N. and D. (2016). Personal Interview. Leavenworth KS, USA: (personal communication).

Kass, M. R. E., Carlin, P. B. P., Gelman, A., & Neal, R. M. (1997). Markov Chain Monte Carlo in Practice : A Roundtable Discussion, 1–26.

Kealey, C. G., Macneil, M. D., Tess, M. W., Geary, T. W., & Bellows, R. A. (2006). Genetic parameter estimates for scrotal circumference and semen characteristics of Line 1 Hereford bulls. *Journal of Animal Science*, *84*, 283–290.

Knights, S. A., Baker, R. L., Gianola, D., & Gibb, J. B. (1984). Estimates of heritabilities and of genetic and phenotypic correlations among growth and reproductive traits in yearling Angus bulls. *Journal of Animal Science*, *58*(4), 887–893.

Kriese, L. A., Bertrand, J. K., & Benyshek, L. L. (1991). Age adjustment factors, heritabilities and genetic correlations for scrotal circumference and related growth traits in Hereford and Brangus bulls. *Journal of Animal Science*, *69*(2), 478–489.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models*. (R. T. Hercher, Ed.) (5th ed.). New York, NY: McGraw-Hill - Irwin.

Land, R. B. (1973). The expression of female sex-limited characters in the male. *Nature*, *241*, 208–209.

Laster, D. B., Glimp, H. A., & Gregory, K. E. (1972). Age and weight at puberty and conception in different breed and breed crosses of beef heifers. *Journal of Animal Science*, *34*, 1031.

Long, P. (2009). Cashing in on the black market. *The Whiteface: Journal of the American Hereford Association*.

Ludwig, C. (2012). The value of Black Hereford registration. *Black Hereford Journal*, *1*(1), 10–13.

Lunstra, D. D. (1982). *Testicular development and onset of puberty in beef bulls*. Clay Center NE, USA.

Lunstra, D. D., Gregory, K. E., & Cundiff, L. V. (1988). Heritability estimates and adjustment factors for the effects of bull age and age of dam on yearling testicular size in breeds of bulls. *Theriogenology*, *30*(1), 127–136.

Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland MA, USA: Sinauer Associates, Inc.

Maniatis, G., Demiris, N., Kranis, A., Banos, G., & Kominakis, A. (2015). Comparison of inference methods of genetic parameters with an application to body weight in broilers. *Archiv Tierzucht*, *58*(2), 277–286.

Martin-Collado, D., Byrne, T. J., Visser, B., & Amer, P. R. (2016). An evaluation of alternative selection indexes for a non-linear profit trait approaching its economic optimum. *Journal of Animal Breeding and Genetics*, *133*, 476–484.

Martinez-Velázquez, G., Gregory, K. E., Bennet, G. L., & Van Vleck, L. D. (2003). Genetic relationships between scrotal circumference and female reproductive traits. *Journal of Animal Science*, *81*, 395–401.

McAllister, C. M., Speidel, S. E., Crews Jr, D. H., & Enns, R. M. (2011). Genetic parameters for intramuscular fat percentage, marbling score, scrotal circumference, and heifer pregnancy in Red Angus cattle. *Journal of Animal Science*, *89*, 2068–2072.

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2015). Manual for BLUPF90 family of programs, 125. Retrieved from http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf.

Morris, C. A., Baker, R. L., & Cullen, N. G. (1992). Genetic correlations between pubertal traits in bulls and heifers, *31*, 221–234.

Moser, D. W., Bertrand, J. K., Benyshek, L. L., McCann, M. A., & Kiser, T. E. (1996). Effects of Selection for Scrotal Circumference in Limousin Bulls on Reproductive and Growth Traits of Progeny. *Journal of Animal Science*, *74*, 2052–2057.

Mrode, R. A. (1996). *Linear models for the prediction of animal breeding values* (1st ed.). Cambridge MA, USA: CABI Publishing.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

Neely, J. D., Johnson, B. H., Dillard, E. U., & Robison, O. W. (1982). Genetic Parameters for testes size and sperm number in Hereford bulls. *Journal of Animal Science*, *55*(5), 1033–1040.

Nelsen, T. C., Long, C. R., & Cartwright, T. C. (1982). Postinflection Growth in Straightbred and Crossbred Cattle. II. Relationships among Weight, Height and Pubertal Characters. *Journal of Animal Science*, *55*(2), 293–304.

Nicholas, F. (2010). *Introduction to veterinary genetics* (3rd ed.). Ames IA, USA: Blackwell Publishing Ltd.

Nunez-dominquez, R., Cundiff, L. V, Dickerson, G. E., Gregory, K. E., & Koch, R. M. (1985). Effects of Managing Heifers to Calve First at Two vs Three Years of Age on Longevity and Lifetime Production of Beef Cows. *U.S. Meat Animal Research Center*, (42), 33–35.

Olshausen, B. A. (2004). *Bayesian probability theory*. (Berkeley Redwood center, Ed.).

Olson, T. A. (1999). Genetics of Colour Variation. In Fries & Ruvinsky (Eds.), *The genetics of cattle* (pp. 33–54). Wallingford Oxfordshire, UK: CABI International.

Patterson, D. J., Perry, R. C., Kiracofe, G. H., Bellows, R. A., Staigmiller, R. B., & Corah, L. R. (1992). Management considerations in heifer development and puberty. *Journal of Animal Science*, *70*(12), 4018–4035.

Perry, G. (2004). *The bovine estrous cycle*. South Dakota State University.

Phillips, C. J. C. (2010). Principles of cattle production. Wallingford Oxfordshire, UK: CABI International.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, *6*, 7–11.

Salverson, R., & Perry, G. (2005). *Understanding estrus sychronization of cattle*.

Schaeffer, L. (2000). *Restricted Maximum Likelihood and Bayesian Estimation*. University of Guelph, Animal Biosciences.

Schaeffer, L. (2010). *Linear Models and Animal Breeding*. *Course*. Guelph, ON: Centre for Genetic Improvement of Livestock Department of Animal and Poultry Science University of Guelph Guelph.

Schmutz, S. (2016). Genetics of Coat Color Patterns in Cattle. Saskatoon, Canada: University of Saskatchewan.

Simm, G. (1998). Genetic improvement of cattle and sheep. Ipswich, UK: Farming Press.

Smith, B. A., Brinks, J. S., & Richardson, G. V. (1989a). Estimation of genetic parameters among reproductive and growth traits in yearling heifers. *Journal of Animal Science*, *67*, 2886–2891.

Smith, B. A., Brinks, J. S., & Richardson, G. V. (1989b). Relationships of sire scrotal circumference to offspring reproduction and growth. *Journal of Animal Science*, *67*, 2881–2885.

Splan, R. K., Cundiff, L. V, & Van Vleck, L. D. (1998). Genetic Parameters for Sex-Specific Traits in Beef Cattle. *Journal of Animal Science*, *76*, 2272–2278.

Tedeschi, L. O. (2006). Assessment of the adequacy of mathematical models. *Agricultural Systems*, *89*(2–3), 225–247.

Toelle, V. D., & Robison, O. W. (1985). Estimates of genetic correlations between testicular measurements and female reproductive traits in cattle. *Journal of Animal Science*, *60*(1), 89–100.

Torres-Vázquez, J. A., & Spangler, M. L. (2016). Genetic parameters for docility, weaning weight, yearling weight, and intramuscular fat percentage in Hereford cattle. *Journal of Animal Science*, *94*(1), 21–27.

Wathes, D. C., Pollott, G. E., Johnson, K. F., Richardson, H., & Cooke, J. S. (2014). Heifer fertility and carry over consequences for life time production in dairy and beef cattle. *Animal*, *8*(1), 91–104.

Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, *22*(22), 3527–3541.

Contours used to create figure 2.1 on page 5 originate from: www.pinterest.com/pin/174021973072877966/ www.123rf.com/stock-photo/animal_moo.html and www.cattle.com/blog/images/Coloring_Pages/Brangus.pdf

Cattle pictures in figure 3.4 on page 21 originate from Joe Hoagland,
www.solutiongenetics.com/blackangus.html and www.toppherefords.com/hereford-herd-bulls/

Vulgarizing summary

The Black Hereford breed is a beef breed created in the 1990s in the USA. During the first years of the breeding program, emphasis has been on the improvement of production and beef quality traits. The last years, breeders want to extend their focus to reproductive traits. One of these traits is the moment at first calving in heifers. In ideal commercial conditions, a heifer calves at an age of two years old. However, in many herds the average age is higher.

One way of reducing the age at first calving, is by selecting for a correlated trait. One of these correlated traits is the yearling scrotal circumference in bulls. In previous breeds, it is shown that selection for a larger yearling scrotal circumference in bulls decreased the age of the onset of puberty in daughters. This made it possible for the age at first calving to decrease.

Measurements on 2532 Black Herefords are included in this study: 894 on yearling scrotal, 2350 on birth weight, 2077 on weaning weight and 1434 on yearling weight. The genetic parameters are estimated using an animal model. This animal model takes both the observations on the animals and, if available, the observations on their relatives (e.g. progeny and parents) into account.

The heritability for yearling scrotal circumference, birth-, weaning- and yearling weight are estimated as moderate to high (0.49, 0.43, 0.47 and 0.45, respectively). The heritability is an estimate to what extent the observed variances can be explained by genetics. A moderate to high heritability (> 0.35) indicates that selection for this trait is possible. Next, a low genetic correlation between yearling scrotal circumference and birth weight is observed (0.22) and a moderate genetic correlation between yearling scrotal circumference and weaning weight (0.54), and between yearling scrotal circumference and yearling weight (0.54). The genetic correlation between weaning- and yearling weight is estimated as high (0.89). A high genetic correlation indicates that when breeders select for the first trait, they automatically will influence the second trait as well. The data of age at first calving were less consistent than the above mentioned weight measurements. This results in the inability to correctly estimate the genetic correlation between yearling scrotal circumference and age at first calving. Breeding values for yearling scrotal circumference, birth-, weaning- and yearling weight are estimated.

This study concludes that the results, previously found in other beef breeds, concerning the heritability of yearling scrotal circumference and the genetic correlations between yearling scrotal circumference and weight traits are also present in the Black Hereford breed.

Appendices

Appendix 1 – The J&N Ranch management

The weight- and YSC data for this research originate from the J&N Ranch, run by Joe and Norma Hoagland. This ranch, located near Leavenworth, Kansas, is the birthplace of the Black Hereford breed. J&N Ranch owns the F. Morgan feedyard, the Decker Farm and the Gunbarrel Ranch in Eskridge, Kansas.

During the grazing season, from the end of April until October, the cattle grazes at the Gunbarrel ranch, which is located in the North of the Flint Hills, Eskridge, Kansas. During the winter period, from October until April, the cattle is fed at the F. Morgan feedyard. At the feedyard, the diet consists of chopped hay, corn, corn silage, distillers grains and additive cow minerals. The postweaning average daily gain during grazing in summer reaches ± 1.1 kg/day. At the feedlot, average daily gain can reach ±1.6 kg/day (Hoagland, 2016).

The breeding season starts every year on May 10th and lasts until July 10th (± 60 days). At the onset of the breeding season all heifers are estrous synchronized. By using hormones, like prostaglandins, progestins and gonadotropins, the heifers will start their next estrus cycle together (Salverson & Perry, 2005). The heifers are artificially inseminated, using semen of one of their own breeding bulls. The insemination is performed at the beginning of the breeding season and afterwards the heifers are brought to the breeding pastures and exposed to a cleanup bull. The cleanup bull will mate with open heifers. Older cows are bred naturally by a chosen breeding bull and are put in single- or multi-sire pastures. When born, the calves are tested for paternity by a DNA-test (Hoagland, 2016).

After a gestation period of ± 283 days, the calving season starts at February 15th and will last until April 15th. Almost 70% of the calves are born in the first 31 days. 97 % of all cows and heifers calve without any assistance. Weaning starts at September 1st and is usually completed by September 15th. By then, the calves are on average 200 days old (Hoagland, 2016).

At the J&N Ranch, every year over 200 cows get bred and over 50 of them are heifers. They own a dozen older bulls. Annually over 150 breeding bulls of various ages are sold to other ranches. They also own a few hundred commercial, non-registered, cattle. In total, the J&N Ranch, with its feedyard and the pasture ranch in Eskridge, has a capacity of over 1500 animals (Hoagland, 2016).

Appendix 2A – BLUP and the MME

As given in chapter 2.3.3, mixed linear models have the following general form (see equation 2.5):

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{e}$$

Consider the expectations (E) of the component vectors of the mixed model (Mrode, 1996):

$$\boldsymbol{E}(\boldsymbol{Y}) = \boldsymbol{X\beta}, \boldsymbol{E}(\boldsymbol{a}) = \boldsymbol{0} \text{ and } \boldsymbol{E}(\boldsymbol{e}) = \boldsymbol{0}$$

The variances are given as following:

$$\boldsymbol{var}(\boldsymbol{e}) = \boldsymbol{I}\boldsymbol{\sigma}_e^2 = \boldsymbol{R}, \boldsymbol{var}(\boldsymbol{a}) = \boldsymbol{A}\boldsymbol{\sigma}_a^2 = \boldsymbol{G} \text{ and } \boldsymbol{cov}(\boldsymbol{a}, \boldsymbol{e}) = \boldsymbol{cov}(\boldsymbol{e}, \boldsymbol{a}) = \boldsymbol{0}$$

where **I** is a n x n identity matrix, σ_e^2 is the residual error variance, σ_a^2 is the additive genetic variance and **A** is the relationship matrix.

Furthermore, it can be calculated that (Mrode, 1996):

$$\begin{aligned}
cov(\boldsymbol{Y}, \boldsymbol{a}) &= cov(\boldsymbol{Za} + \boldsymbol{e}, \boldsymbol{a}) \\
&= cov(\boldsymbol{Za}, \boldsymbol{a}) + cov(\boldsymbol{e}, \boldsymbol{a}) \\
&= \boldsymbol{Z}\, cov(\boldsymbol{a}, \boldsymbol{a}) \\
&= \boldsymbol{Z}\, var(\boldsymbol{a}) \\
&= \boldsymbol{ZG} \\
cov(\boldsymbol{Y}, \boldsymbol{e}) &= cov(\boldsymbol{Za} + \boldsymbol{e}, \boldsymbol{e}) \\
&= cov(\boldsymbol{Za}, \boldsymbol{e}) + cov(\boldsymbol{e}, \boldsymbol{e}) \\
&= \boldsymbol{Z}\, cov(\boldsymbol{a}, \boldsymbol{e}) + cov(\boldsymbol{e}, \boldsymbol{e}) \\
&= cov(\boldsymbol{e}, \boldsymbol{e}) \\
&= var(\boldsymbol{e}, \boldsymbol{e}) \\
&= \boldsymbol{R} \\
var(\boldsymbol{Y}) &= \boldsymbol{V} \\
&= var(\boldsymbol{Za} + \boldsymbol{e}) \\
&= \boldsymbol{Z}\, var(\boldsymbol{a})\, \boldsymbol{Z}' + var(\boldsymbol{e}) + cov(\boldsymbol{Za}, \boldsymbol{e}) + cov(\boldsymbol{e}, \boldsymbol{Za}) \\
&= \boldsymbol{ZGZ}' + \boldsymbol{R} + \boldsymbol{Z}\, cov(\boldsymbol{a}, \boldsymbol{e}) + cov(\boldsymbol{e}, \boldsymbol{a})\boldsymbol{Z}'
\end{aligned}$$

Since $cov(\boldsymbol{a}, \boldsymbol{e}) = cov(\boldsymbol{e}, \boldsymbol{a}) = 0$.

Then:

$$\boldsymbol{V} = \boldsymbol{ZGZ}' + \boldsymbol{R}$$

This expression indicates that the observed variance can be attributed to two separate terms. The first term accounts for the contribution from random genetic effects, while the second accounts for the variance due to residual effects.

Next, $\boldsymbol{a}$ will be predicted and $\boldsymbol{\beta}$ will be estimated (this calculation is derived from Mrode, 1996).

Call $\boldsymbol{k}'\boldsymbol{\beta}+\boldsymbol{a}$ the predictant. Suppose that $\boldsymbol{k}'\boldsymbol{\beta}+\boldsymbol{a}$ is a linear function of $\boldsymbol{\beta}$ and $\boldsymbol{a}$. $\boldsymbol{k}'\boldsymbol{\beta}+\boldsymbol{a}$ is predictable using a linear function of $\boldsymbol{Y}$, say $\boldsymbol{L}'\boldsymbol{Y}$, and $\boldsymbol{k}'\boldsymbol{\beta}$ is estimable. The predictor $\boldsymbol{L}'\boldsymbol{Y}$ can be chosen such that $E(\boldsymbol{L}'\boldsymbol{Y}) = E(\boldsymbol{k}'\boldsymbol{\beta}+\boldsymbol{a})$ is unbiased (i.e. its expected value is equal to the expected value of the predictant) and the PEV is minimized (Mrode, 1996). The PEV can be calculated as (Mrode, 1996):

$$\begin{aligned} PEV &= var(\boldsymbol{a}-\widehat{\boldsymbol{a}}) \\ &= var(\boldsymbol{L}'\boldsymbol{Y}-\boldsymbol{k}'\boldsymbol{\beta}+\boldsymbol{a}) \\ &= var(\boldsymbol{L}'\boldsymbol{Y}-\boldsymbol{a}) \\ &= \boldsymbol{L}'\,var(\boldsymbol{Y})\,\boldsymbol{L}+var(\boldsymbol{a})-\boldsymbol{L}'cov(\boldsymbol{Y},\boldsymbol{a})-cov(\boldsymbol{a},\boldsymbol{Y})\boldsymbol{L} \\ &= \boldsymbol{L}'\boldsymbol{V}\boldsymbol{L}+\boldsymbol{G}-\boldsymbol{L}'\boldsymbol{V}\boldsymbol{G}-\boldsymbol{Z}\boldsymbol{G}'\boldsymbol{L} \end{aligned}$$

Minimizing PEV subject to $E(\boldsymbol{L}'\boldsymbol{Y}) = E(\boldsymbol{k}'\boldsymbol{\beta}+\boldsymbol{a})$ and solving gives (Henderson, 1984):

$$\boldsymbol{L}'\boldsymbol{Y} = \boldsymbol{k}'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y}-\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{V}^{-1}\left(\boldsymbol{Y}-\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y}\right)$$

The BLUP of **a** equals to $\widehat{\boldsymbol{a}} = \boldsymbol{G}\boldsymbol{Z}'\boldsymbol{V}^{-1}\left(\boldsymbol{Y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)$**.** The BLUE (best linear unbiased estimator) of **β** equals $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y}$. BLUE is an estimator of the estimable functions of fixed effects that has minimum sampling variance, is unbiased and is based on a linear function of the data. This way, the predictor can be written as (Mrode, 1996):

$$\boldsymbol{L}'\boldsymbol{Y} = \boldsymbol{k}'\widehat{\boldsymbol{\beta}}+\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}})$$

Which is the BLUP of **k' β + a**.

The solutions for **a** and **β** require the computation of $\mathbf{V}^{-1}$. This $\mathbf{V}^{-1}$ matrix is often so large that its inverse is almost impossible to calculate. To solve this problem, Henderson calculated in 1950 a set of equations which estimate **β** and predict **a** simultaneously, without the need for computing $\mathbf{V}^{-1}$. These equations are called the mixed-model equations (MME).

The MME of the general equation are, assuming non-singularity of **R** and **G** (Mrode, 1996):

$$\begin{bmatrix} \boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z}+\boldsymbol{G}^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{Y} \\ \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Y} \end{bmatrix}$$

The advantage of using the MME over the general form is that neither **V** nor its inverse is required. Although **R** has the same dimensionality as **V**, calculating the inverse is more easy since it is an identity matrix. Furthermore, **G** is often diagonal and thus is **G**$^{-1}$ trivial to obtain. **Z'R**$^{-1}$**Z + G**$^{-1}$ is either a diagonal matrix or has a large diagonal submatrix. **R**$^{-1}$, being an identity matrix, can be factorized from both sides of the MME to give (Mrode, 1996):

$$\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{Z}+\boldsymbol{A}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{Y} \\ \boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Y} \end{bmatrix}$$

where $\alpha = \sigma_e^2/\sigma_a^2$ or $(1-h^2)/h^2$.

The solutions to the MME give the BLUE of **k' β** and the BLUP of **a**.

Appendix 2B – BLUP and accuracy values

The calculation of the accuracy value (r) requires the diagonal elements of the inverse of the MME. For simplification, the coefficient matrix of the MME can be represented as (Mrode, 1996):

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \boldsymbol{A}^{-1}\alpha \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix}$$

The left hand side of this equation can be inverted as:

$$\begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{C}^{11} & \boldsymbol{C}^{12} \\ \boldsymbol{C}^{21} & \boldsymbol{C}^{22} \end{bmatrix}$$

Then (Mrode, 1996):

$$\begin{aligned} PEV &= var(a - \hat{a}) \\ &= \boldsymbol{C}^{22}\sigma_e^2 \\ &= (1 - r^2)\sigma_a^2 \end{aligned}$$

where *r²* is the squared correlation between the true and estimated breeding values, called the reliability.

For animal *i* (Mrode, 1996):

$$d_i\sigma_e^2 = (1 - r_i^2)\sigma_a^2$$

where d_i is the *i*-th diagonal element of $\boldsymbol{C_{22}}$ and r_i the accuracy value of the prediction for animal *i*. r_i can be calculated as:

$$\frac{d_i\sigma_e^2}{\sigma_a^2} = 1 - r_i^2$$

and thus:

$$r_i = \sqrt{1 - \frac{d_i\sigma_e^2}{\sigma_a^2}}$$

Appendix 3 – Bayesian statistics: Gibbs sampling

This section is for the largest part an excerpt from the course notes 'Linear models and animal breeding' by Larry Schaeffer, University of Guelph (course given: June 2010 in Norway) (Schaeffer, 2010).

The description of Bayesian statistics is based on the vector of random variables $\boldsymbol{\theta}$, the data vector, $\boldsymbol{Y}$, and the mixed linear animal model:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{e}$$

As given in chapter 2.3.6, Bayesian statistics is based on the Bayes' theorem, where $p(\boldsymbol{\theta}|\boldsymbol{Y})$ is the posterior probability function of $\boldsymbol{\theta}$:

$$p(\boldsymbol{\theta}|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\boldsymbol{Y})}$$

where $p(\boldsymbol{Y}|\boldsymbol{\theta})$ is the probability of the observed data $\boldsymbol{Y}$, called the likelihood function, $p(\boldsymbol{\theta})$ is called the prior probability and $p(\boldsymbol{Y})$ is obtained by integrating $p(\boldsymbol{Y}|\boldsymbol{\theta})\, p(\boldsymbol{\theta})$ over all $\boldsymbol{\theta}$.

In terms of the animal model, $\boldsymbol{\theta}$ includes all unknown parameters ($\boldsymbol{\beta}$, $\boldsymbol{a}$, σ_a^2 and σ_e^2).

The conditional distribution of $\boldsymbol{Y}$, given $\boldsymbol{\theta}$, is equal to:

$$\boldsymbol{Y} \mid \boldsymbol{\beta}, \boldsymbol{a}, \sigma_a^2, \sigma_e^2 \sim N\left(\boldsymbol{X}\text{ß} + \boldsymbol{Za}, \boldsymbol{I}\boldsymbol{\sigma_e^2}\right) \text{ and}$$

$$p(\boldsymbol{Y} \mid \boldsymbol{\beta}, \boldsymbol{a}, \sigma_a^2, \sigma_e^2) \propto (\sigma_e^2)^{\left(-\frac{n}{2}\right)} exp\left[-(\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Za})' \frac{(\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Za})}{2\sigma_e^2}\right]$$

where *n* is equal to the number of observations.

To start, prior distributions need to be assigned to $\boldsymbol{\theta}$. To do so, every individual term of $\boldsymbol{\theta}$ has to receive a prior distribution. Then all these priors are multiplied to each other and to the conditional distribution of $\boldsymbol{Y}$.

These prior distributions are given as following:

- For the fixed effects vector, $\boldsymbol{\beta}$, there is little prior knowledge about the values that elements in that vector might have. This is represented by assuming that

$$p(\boldsymbol{\beta}) \propto constant$$

- For $\boldsymbol{a}$, the vector of additive genetic values, quantitative genetics theory suggests that it follows a Normal distribution:

$$\boldsymbol{a} \mid \boldsymbol{A}, \sigma_a^2 \sim N(\boldsymbol{0}, \boldsymbol{A}\sigma_a^2) \text{ and}$$

$$p(\boldsymbol{a}) \;\propto\; (\sigma_a^2)^{\left(-\frac{q}{2}\right)} \exp\left[-\boldsymbol{a}'\boldsymbol{A}^{-1}\frac{\boldsymbol{a}}{2\sigma_a^2}\right],$$

  where q is the length of $\boldsymbol{a}$. A natural estimator of σ_a^2 is $\frac{\boldsymbol{a}'\boldsymbol{A}^{-1}\boldsymbol{a}}{\boldsymbol{q}}$.

  Call this expression S_a^2. It follows a scaled chi-square distribution:

$$S_a^2 \sim \frac{\chi_q^2 \sigma_a^2}{q}.$$

  Multiplying both sides by q and dividing by χ_q^2 gives:

$$\sigma_a^2 \sim \frac{q S_a^2}{\chi_q^2},$$

  which is a scaled inverted Chi-square distribution and can be written as:

$$p(\sigma_a^2 | v_a, S_a^2) \propto (\sigma_a^2)^{-(\frac{v_a}{2}+1)} exp\left(-\frac{v_a}{2}\frac{S_a^2}{\sigma_a^2}\right),$$

  where v_a and S_a^2 are parameters of the prior distribution, with S_a^2 being a prior guess about the value of σ_a^2 and v_a being the degrees of belief in that prior value. Usually q is much larger than v_a and therefore, the data provide nearly all of the information about σ_a^2.

- Similarly, for the residual variance (σ_e^2), the following expression can be given:

$$p(\sigma_e^2 | v_e, S_e^2) \propto (\sigma_e^2)^{-\left(\frac{v_e}{2}+1\right)} exp\left(-\frac{v_e}{2}\frac{S_e^2}{\sigma_e^2}\right).$$

The joint posterior distribution can then be written as:

$$p(\boldsymbol{\beta}, \boldsymbol{a}, \sigma_a^2, \sigma_e^2 | Y) \propto p(\boldsymbol{\beta})\ p(\boldsymbol{a}|\sigma_a^2)\ p(\sigma_a^2)\ p(\sigma_e^2)\ p(\boldsymbol{Y}|\boldsymbol{\beta}, \boldsymbol{a}, \sigma_a^2, \sigma_e^2)$$

or:

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{a}, \sigma_a^2, \sigma_e^2 | \boldsymbol{Y}) & \\ &\propto\ (\sigma_e^2)^{-\left(\frac{n+v_e}{2}+1\right)} \exp\left[-\frac{1}{2\sigma_e^2}((\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Za})'(\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Za}) \right. \\ &\left. +\ v_e S_e^2)\right]\ (\sigma_a^2)^{-\left(\frac{q+v_a}{2}+1\right)} \exp\left[-\frac{1}{2\sigma_a^2}(\boldsymbol{a}'\boldsymbol{A}^{-1}\boldsymbol{a} + v_a S_a^2)\right]. \end{aligned}$$

In order to implement Gibbs sampling, all of the fully conditional posterior distributions (one for each component of $\boldsymbol{\theta}$) need to be derived from the above joint posterior distribution.

Say that:

$$\boldsymbol{W} = (\boldsymbol{X}\,\boldsymbol{Z})$$
$$\boldsymbol{\gamma}' = (\boldsymbol{\beta}'\boldsymbol{a}')$$
$$\Sigma = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}^{-1}k \end{bmatrix}$$
$$\boldsymbol{C} = \boldsymbol{W}'\boldsymbol{W} + \Sigma$$
$$\boldsymbol{C}\hat{\boldsymbol{\gamma}} = \boldsymbol{W}'\boldsymbol{y}$$

This last statement, is equal to the MME.

Now we adopt a new notation, let:

$$\gamma' = (\gamma_i\ \gamma_{-i}')$$

where γ_i is a scalar representing just one element of the vector $\boldsymbol{\gamma}$ (gamma), and γ_{-i} is a vector representing all of the other vector elements except γ_i. Similarly, $\boldsymbol{C}$ and $\boldsymbol{W}$ can be partitioned in the same way as:

$$\boldsymbol{W}' = (\boldsymbol{W}_i\ \boldsymbol{W}_{-i})'$$
$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{i,i} & \boldsymbol{C}_{i,-i} \\ \boldsymbol{C}_{-i,i} & \boldsymbol{C}_{-i,-\boldsymbol{i}} \end{bmatrix}.$$

In general terms, the conditional posterior distribution of γ is equal to:

$$\gamma_i|\gamma_{-i},\sigma_a^2,\sigma_e^2,\boldsymbol{y} \sim N(\widehat{\gamma_i},C_{i,i}^{-1}\sigma_e^2)$$

where

$$C_{i,i}\hat{\gamma}_i = \left(\boldsymbol{W}_i'\boldsymbol{y} - \boldsymbol{C}_{i,-i}\gamma_{-\boldsymbol{i}}\right).$$

Then

$$\beta_i|\beta_{-i},\boldsymbol{a},\sigma_a^2,\sigma_e^2,\boldsymbol{y} \sim N\left(\widehat{\beta_i},C_{i,i}^{-1}\sigma_e^2\right)$$

for

$$C_{i,i} = \boldsymbol{x}_i'\boldsymbol{x}_i.$$

Also,

$$a_i|\boldsymbol{\beta},\boldsymbol{a}_{-i},\sigma_a^2,\sigma_e^2,\boldsymbol{y} \sim N\left(\widehat{a_i},C_{i,i}^{-1}\sigma_e^2\right)$$

where

$$C_{i,i} = \left(\boldsymbol{z}_i'\boldsymbol{z}_i + A^{i,i}k\right)$$

for $k = \frac{\sigma_a^2}{\sigma_e^2}$. The conditional posterior distributions for the variances are

$$\sigma_a^2|\boldsymbol{\beta},\boldsymbol{a},\sigma_e^2,\boldsymbol{y} \sim \tilde{v}_a\tilde{S}_a^2\chi_{\tilde{v}_a}^{-2}$$

for $\tilde{v}_a = q + v_a$, and $\tilde{S}_a^2 = \frac{\boldsymbol{a}'\boldsymbol{A}^{-1}\boldsymbol{a}+v_aS_a^2}{\tilde{v}_a}$, and

$$\sigma_e^2|\boldsymbol{\beta},\boldsymbol{a},\sigma_a^2,\boldsymbol{y} \sim \tilde{v}_e\tilde{S}_e^2\chi_{\tilde{v}_e}^{-2}$$

for $\tilde{v}_e = N + v_e$, and $\tilde{S}_e^2 = \frac{\boldsymbol{e}'\boldsymbol{e}+v_eS_e^2}{\tilde{v}_e}$, and $\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{Za}$.

Appendix 4 – Parameter file for renumf90

This is an example of a parameter file, used as input file for the renumf90 program. This file was used to estimate model 3.

```
# Black Hereford - Roel Meyermans - Model 3
# Parameter file for program renf90; it is translated to parameter
# file for BLUPF90 family programs.
#
DATAFILE
data.txt
TRAITS
2 3 4 6                    #YSC BW WW YW
FIELDS_PASSED

WEIGHT(S)

RESIDUAL_VARIANCE
10 1 1 1
1 100 1 1
1 1 10000 1
1 1 1 10000
EFFECT
1 1 1 1 cross alpha        #additive genetic effect
RANDOM
animal
FILE
pedigree.txt
FILE_POS
1 2 3 0 4
PED_DEPTH
12
EFFECT
12 12 12 12 cross alpha    #birth year
EFFECT
0 10 10 0 cross alpha      #ET
EFFECT
0 11 11 11 cross alpha     #Sex
EFFECT
8 8 8 8 cross alpha        #Age of dam
EFFECT
0 0 5 0 cov                #WA
EFFECT
7 0 0 7 cov                #YA
OPTION sol se
OPTION residual
OPTION EM-REML 5
```

Appendix 5 – Results of the linear regression of the four traits

Chapter 8.4 presents the linear regression results of the four traits. This appendix shows the ANOVA results of these four models (table A5.1) (Kutner *et al.*, 2005). Table A5.2 shows a statistic summary of these models. In these models, the age at measurement was scaled as given in chapter 8.5.1.

Table A5.1 Overview of the linear regression results for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW)

YSC						
		Df	Sum of squares	Mean squares	F-value	Pr (>F)
	Year.Age	1	231.4	231.4	42.5	< 0.001
	Age of Dam	5	190.8	38.2	7.0	< 0.001
	Birth year	14	1367.9	97.7	17.9	< 0.001
	Sire	58	413.5	7.1	1.3	0.068
	Residuals	628	3418.8	5.4		
BW						
		Df	Sum of squares	Mean squares	F-value	Pr (>F)
	Age of Dam	5	3113.8	622.8	52.5	< 0.001
	Birth year	19	2936.7	154.0	13.0	< 0.001
	Sex	1	2869.8	2869.8	241.9	< 0.001
	ET	1	113.0	113.0	9.5	0.002
	Sire	79	3578.4	45.3	3.8	< 0.001
	Residuals	2130	25274.0	11.9		
WW						
		Df	Sum of squares	Mean squares	F-value	Pr (>F)
	Wean.Age	1	354630	354630	489.0	< 0.001
	Age of Dam	5	394068	78814	108.7	< 0.001
	Birth year	17	236570	13916	19.2	< 0.001
	Sex	1	195078	195078	269.0	< 0.001
	ET	1	43572	43572	60.1	< 0.001
	Sire	68	180627	2656	3.7	< 0.001
	Residuals	1957	1419241	725		
YW						
		Df	Sum of squares	Mean squares	F-value	Pr (>F)
	Year.Age	1	571580	571580	416.2	< 0.001
	Age of Dam	5	331680	66336	48.3	< 0.001
	Birth year	15	1253643	83576	60.9	< 0.001
	Sex	1	1076765	1076765	784.0	< 0.001
	Sire	62	256396	4135	3.0	< 0.001
	Residuals	1326	1821127	1373		

Table A5.2 Summary of the number of observations, adjusted R² (Kutner et al., 2005) and root mean squared error (RMSE) (Kutner et al., 2005) for the univariate linear regression models of yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW)

	YSC	BW	WW	YW
Number of observations	707	2236	2053	1411
Adjusted R²	0.316	0.300	0.474	0.635
RMSE	2.334	3.445	26.930	37.060

Appendix 6 – (Co)variance components obtained using REML

The following table (A6.1) contains the REML (co)variance estimates of the additive genetic and the residual effect of the five models. The additive genetic (co)variances be found in the first diagonal matrix. The second diagonal matrix contains the (co)variances of the residual effect. The standard errors for every estimate is given between parentheses.

Table A6.1 Variance and covariance estimates for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) estimated by the five models. The first diagonal matrix gives the additive genetic (co)variances for each model, the second diagonal matrix gives the residual (co)variances for the model. The value between parantheses gives the standard error

Model 1					
		YSC	BW	WW	YW
	YSC	2.24 (0.60)	0.75 (0.52)	14.87 (4.65)	-2.72 (4.89)
	BW		5.77 (0.77)	16.71 (4.76)	4.08 (6.70)
	WW			403.13 (52.76)	8.27 (56.42)
	YW				76.78 (69.19)
	YSC	3.71 (0.46)	0.88 (0.40)	17.19 (3.44)	7.40 (5.41)
	BW		7.86 (0.50)	14.92 (3.05)	1.35 (6.70)
	WW			461.0 (33.18)	102.11 (56.43)
	YW				2692.60 (69.19)
Model 2					
		YSC	BW	WW	YW
	YSC	2.93 (0.63)	0.64 (0.54)	20.89 (4.97)	27.91 (7.26)
	BW		5.76 (0.77)	15.82 (4.90)	29.31 (7.40)
	WW			444.48 (57.11)	535.33 (75.33)
	YW				799.54 (125.99)
	YSC	4.00 (0.48)	0.65 (0.41)	19.07 (3.60)	38.31 (5.45)
	BW		7.87 (0.77)	13.30 (6314)	12.92 (5.00)
	WW			509.53 (57.11)	474.88 (48.86)
	YW				1161.50 (88.67)
Model 3					
		YSC	BW	WW	YW
	YSC	3.41 (0.66)	0.99 (0.56)	20.10 (4.89)	30.42 (7.24)
	BW		5.80 (0.77)	17.11 (4.74)	33.38 (7.26)
	WW			405.12 (52.81)	501.65 (69.93)
	YW				784.12 (118.41)
	YSC	3.60 (0.48)	0.69 (0.41)	16.76 (3.48)	29.47 (5.15)
	BW		7.86 (0.49)	14.73 (3.03)	15.49 (4.75)
	WW			464.58 (32.98)	413.50 (44.22)
	YW				941.67 (79.15)

Model 4				
	YSC	BW	WW	YW
YSC	3.29 (0.65)	1.01 (0.56)	19.26 (4.84)	29.46 (7.15)
BW		5.85 (0.77)	18.15 (4.74)	34.99 (7.27)
WW			401.91 (52.27)	495.58 (69.40)
YW				780.08 (117.54)
YSC	3.67 (0.48)	0.67 (0.41)	17.46 (3.47)	30.04 (5.14)
BW		7.82 (0.49)	13.92 (3.02)	14.29 (4.74)
WW			469.44 (32.97)	421.57 (44.12)
YW				948.57 (78.89)

Model 5				
	YSC	BW	WW	YW
YSC	3.36 (0.66)	0.98 (0.56)	20.05 (4.87)	30.33 (7.21)
BW		5.80 (0.77)	17.10 (4.74)	33.39 (7.26)
WW			405.34 (52.46)	502.16 (69.91)
YW				785.12 (118.39)
YSC	3.64 (0.48)	0.69 (0.41)	16.88 (3.47)	29.65 (5.15)
BW		7.86 (0.49)	14.74 (3.03)	15.49 (4.75)
WW			464.51 (32.97)	413.33 (44.20)
YW				941.28 (79.11)

The variances for the birth year effect, as estimated in model 4, are equal to 3.94, 1.56, 192.21 and 237.06 for YSC, BW, WW and YW, respectively.

Table A6.2 gives the genetic correlations in the upper diagonal matrix, and the residual correlations in the lower diagonal matrix for the five models. The heritability estimates are given on the diagonal and highlighted in grey.

Table A6.2 The estimated genetic correlations (upper diagonal matrix), residual correlations (lower diagonal matrix) and heritabilities (on the diagonal, highlighted in grey) for yearling scrotal circumference (YSC), birth weight (BW), weaning weight (WW) and yearling weight (YW) in the five models

Model 1					
		YSC	BW	WW	YW
	YSC	0.38	0.21	0.49	-0.21
	BW	0.16	0.42	0.34	0.19
	WW	0.41	0.25	0.47	0.05
	YW	0.07	0.09	0.09	0.03
Model 2					
		YSC	BW	WW	YW
	YSC	0.42	0.16	0.57	0.58
	BW	0.12	0.42	0.31	0.43
	WW	0.42	0.21	0.47	0.90
	YW	0.56	0.14	0.62	0.45
Model 3					
		YSC	BW	WW	YW
	YSC	0.49	0.22	0.54	0.59
	BW	0.13	0.43	0.35	0.49
	WW	0.41	0.24	0.47	0.89
	YW	0.51	0.18	0.63	0.45
Model 4					
		YSC	BW	WW	YW
	YSC	0.30	0.23	0.53	0.58
	BW	0.12	0.38	0.37	0.52
	WW	0.42	0.23	0.38	0.88
	YW	0.51	0.17	0.62	0.40
Model 5					
		YSC	BW	WW	YW
	YSC	0.48	0.22	0.54	0.59
	BW	0.13	0.42	0.35	0.49
	WW	0.41	0.24	0.47	0.89
	YW	0.51	0.18	0.63	0.45

Appendix 7 – Parameter estimates

The next four tables (A7.1, A7.2, A7.3 and A7.4) give the parameter estimates (BLUEs) for the five models estimated by BLUP for YSC, BW, WW and YW, respectively.

Yearling scrotal circumference

Table A7.1 Overview of the parameter estimates (in cm) of the five models for yearling scrotal circumference

		Model 1	Model 2	Model 3	Model 4	Model 5
Mean additive genetic effect		0.09	0.21	0.63	0.61	0.60
Birth year effect						
	2001	-	33.07	33.40	34.66	33.44
	2002	35.19	32.43	33.79	33.66	33.90
	2003	34.53	32.86	33.59	33.65	33.71
	2004	43.71	41.09	42.32	41.83	42.58
	2005	38.37	36.18	37.00	36.82	37.12
	2006	36.63	34.30	35.19	35.04	35.30
	2007	36.74	35.28	35.99	35.85	36.05
	2008	37.18	34.10	35.11	35.13	35.25
	2009	37.85	35.10	35.73	35.67	35.86
	2010	38.76	36.30	36.59	36.48	36.73
	2011	38.45	35.72	36.08	36.00	36.22
	2012	36.70	35.13	34.99	34.97	35.19
	2013	37.30	36.48	36.01	35.94	36.15
	2014	35.62	34.88	34.44	35.77	36.15
	2015	35.21	32.98	36.61	35.00	35.31
	2016	37.84	35.64	36.26	37.50	37.97
Observer						
	1	-	-	-	0.00	0.00
	2	-	-	-	1.40	1.57
	3	-	-	-	0.67	0.68
Age of dam						
	2	-1.92	-1.39	-1.11	-0.99	-1.20
	3	-1.32	-0.97	-0.65	-0.51	-0.72
	4	-1.06	-0.80	-0.41	-0.28	-0.49
	5	-0.91	-0.64	-0.24	0.09	-0.32
	6	-1.65	-1.16	-0.74	-0.60	-0.81
	7	-0.72	-0.58	-0.16	-0.02	-0.24
Yearling age as covariate		-	0.0043	0.82	0.82	0.83

Missing values, indicated by a “-“, are not estimated by the model (e.g. caused by a lack of measurements for that effect).

Birth weight

Table A7.2 Overview of the parameter estimates (in kg) of the five models for birth weight

		Model 1	Model 2	Model 3	Model 4	Model 5
Mean additive genetic effect		-0.02	-0.02	0.12	0.14	0.11
Birth year effect						
	1992	-5.33	-5.56	-5.47	-0.46	-5.47
	1993	0.51	0.51	0.53	0.29	0.53
	1997	0.94	0.98	0.92	0.77	0.92
	1998	-2.75	-2.76	-2.76	-0.13	-2.76
	1999	-1.20	-1.26	-1.24	0.73	-1.24
	2000	-0.82	-0.88	-0.89	0.81	-0.89
	2001	-2.37	-2.43	-2.43	-0.47	-2.43
	2002	-5.95	-5.98	-5.97	-3.65	-5.97
	2003	-2.75	-2.79	-2.80	-0.77	-2.80
	2004	-1.01	-1.08	-1.09	0.89	-1.09
	2005	-1.09	-1.15	-1.16	0.76	-1.16
	2006	-3.76	-3.84	-3.87	-1.82	-3.86
	2007	-1.87	-1.93	-1.94	-0.03	-1.94
	2008	-0.28	-0.33	-0.39	1.42	-0.39
	2009	-2.44	-2.51	-2.60	-0.68	-2.60
	2010	-1.25	-1.32	-1.43	0.39	-1.43
	2011	-1.58	-1.68	-1.84	-0.03	-1.84
	2012	-0.74	-0.83	-0.99	0.72	-0.98
	2013	-0.70	-0.81	-0.96	0.77	-0.95
	2014	-0.72	-0.86	-1.01	0.72	-1.00
	2015	-1.93	-2.07	-2.18	-0.37	-2.17
	2016	-1.72	-1.86	-1.92	-0.12	-1.92
ET						
	Yes	38.13	38.11	38.54	36.66	38.54
	No	39.37	39.37	39.30	37.40	39.31
Sex						
	Female	-0.58	-0.57	-0.60	-0.59	-0.60
	Male	1.66	1.67	1.62	1.62	1.62
Age of dam						
	2	-2.27	-2.23	-2.24	-2.27	-2.24
	3	-0.75	-0.73	-0.76	-0.81	-0.76
	4	-0.37	-0.36	-0.38	-0.38	-0.38
	5	0.11	0.13	0.11	0.11	0.11
	6	0.05	0.07	0.07	0.06	0.07
	7	0.24	0.25	0.22	0.23	0.22

Weaning weight

Table A7.3 Overview of the parameter estimates (in kg) of the five models for weaning weight

		Model 1	Model 2	Model 3	Model 4	Model 5
Mean additive genetic effect		-2.79	-1.71	-0.10	0.16	-0.19
Birth year effect						
	1992	-	-	-	226.63	-
	1993	-	-	-	230.58	-
	1997	235.69	215.74	243.46	232.92	243.60
	1998	-	-	-	228.37	-
	1999	-	367.00	351.68	256.12	351.85
	2000	199.28	123.06	196.00	208.59	196.13
	2001	219.71	155.75	219.51	218.24	219.66
	2002	209.44	130.37	207.69	205.50	207.86
	2003	237.68	160.68	235.64	233.48	235.82
	2004	242.19	157.95	239.04	237.68	239.21
	2005	255.14	169.76	251.80	249.03	251.96
	2006	219.16	136.19	215.82	214.61	216.00
	2007	228.18	139.85	224.57	223.19	224.71
	2008	239.41	146.77	234.39	233.08	234.59
	2009	249.71	157.15	244.00	241.96	244.20
	2010	233.27	140.48	226.88	225.07	227.10
	2011	237.88	145.31	230.57	228.81	230.79
	2012	223.66	133.11	216.88	215.32	217.25
	2013	241.98	158.45	236.15	234.01	236.42
	2014	245.82	153.49	239.00	236.87	239.28
	2015	233.66	142.29	227.86	226.64	228.11
	2016	238.46	147.32	233.02	230.91	233.28
ET						
	Yes	-35.27	-25.48	-25.02	-24.96	-25.02
	No	0.00	0.00	0.00	0.00	0.00
Sex						
	Female	12.13	17.48	13.43	13.44	13.43
	Male	31.96	35.40	32.70	32.86	32.70
Age of dam						
	2	-27.39	-37.25	-27.99	-26.77	-28.02
	3	-16.20	-28.19	-17.61	-16.38	-17.65
	4	-5.49	-17.56	-7.00	-5.75	-7.03
	5	-1.93	-13.20	-3.12	-1.75	-3.16
	6	8.81	-3.15	7.97	8.75	7.94
	7	8.83	-2.40	7.23	8.70	7.18
Weaning age as covariate		-	0.46	18.57	18.61	18.58

Missing values, indicated by a “-“, are not estimated by the model (e.g. caused by a lack of measurements for that effect).

Yearling weight

Table A7.4 Overview of the parameter estimates (in kg) of the five models for yearling weight

	Model 1	Model 2	Model 3	Model 4	Model 5
Mean additive genetic effect	-0.21	-1.06	2.07	2.44	1.93
Birth year effect					
1992	-	-	-	420.53	-
1993	-	-	-	421.63	-
1997	321.85	455.77	426.62	420.02	426.85
1998	-	-	-	421.01	-
1999	-	-	-	426.00	-
2000	-	-	-	418.42	-
2001	-	-	-	420.28	-
2002	368.25	337.65	411.78	410.63	412.08
2003	382.83	370.51	426.48	425.19	426.80
2004	344.41	310.80	395.08	394.85	395.39
2005	415.75	376.67	441.44	437.81	441.72
2006	397.96	360.38	426.60	424.81	426.93
2007	393.69	356.61	417.32	415.94	417.54
2008	368.94	328.56	403.97	403.63	404.33
2009	376.26	348.92	417.99	416.36	418.34
2010	397.21	380.94	436.97	434.66	437.35
2011	396.32	357.15	425.47	423.57	425.84
2012	396.18	381.08	432.61	430.40	433.25
2013	404.92	396.45	435.71	432.71	436.18
2014	401.50	399.51	435.36	432.29	435.84
2015	377.33	328.26	395.40	394.73	395.83
2016	412.01	378.98	444.08	440.56	444.50
Sex					
Female	-29.60	-52.41	-69.52	-70.14	-69.54
Male	0.00	0.00	0.00	0.00	0.00
Age of dam					
2	-6.86	-48.01	-38.63	-37.34	-38.71
3	4.57	-42.57	-30.16	-28.79	-30.25
4	6.70	-29.72	-17.63	-16.18	-17.69
5	8.57	-28.43	-17.57	-15.88	-17.65
6	18.43	-19.52	-7.13	-5.73	-7.20
7	0.30	-18.79	-8.21	-6.38	-8.32
Yearling age as covariate	-	0.21	27.23	27.23	27.23

Missing values, indicated by a “-“, are not estimated by the model (e.g. caused by a lack of measurements for that effect).

Appendix 8 – Output of the MCMC algorithm

Table A8.1 gives an overview of the means of the approximated posterior distributions, the 95% the highest posterior density intervals and the effective sample sizes for the effects of the intercept, age of dam, birth year, yearling age and the residuals. The following figure (A8.1) gives an overview of the chain traces (on the left side for each effect) and the approximated posterior distribution (on the right side) for each effect.

Table A8.1 Overview of the posterior means of the approximated posterior density functions (in cm), the 95% posterior density intervals (HPD) and the effective sample sizes for the effects included in the animal model with yearling scrotal circumference as dependent variable

Effect		Posterior mean	Lower – 95% HPD	Upper – 95% HPD	Effective sample size
Additive genetic		2.12	0.95	3.45	1708
Intercept		36.29	35.34	37.18	1825
Age of dam					
	3	0.79	0.16	1.35	1994
	4	0.86	0.26	1.51	1817
	5	1.18	0.47	1.80	1796
	6	0.24	-0.46	0.92	1743
	7	1.11	0.49	1.75	1777
Birth year					
	2002	-3.20	-4.78	-1.60	1776
	2003	-3.81	-5.08	-2.5	2203
	2004	5.08	3.37	6.70	1994
	2005	0.08	-1.06	1.26	1848
	2006	-1.14	-2.05	-0.10	1994
	2007	-0.64	-1.71	0.40	1994
	2008	-0.37	-1.35	0.67	1871
	2009	-0.06	-0.94	0.81	1994
	2010	1.3	0.39	2.26	1994
	2011	0.68	-0.15	1.57	1994
	2012	-1.22	-2.03	-0.43	1994
	2013	-0.26	-1.33	0.66	1994
	2014	-2.03	-2.98	-1.03	1844
	2015	-2.71	-3.55	-2.00	1994
	2016	-	-	-	-
Yearling age as covariate		1.17	0.86	1.48	1776
Residuals		3.66	2.59	4.62	1717

Missing values, indicated by a “-“, are not estimated by the model.

The lower and upper ranges of the 95% highest posterior density interval for the heritability are equal to 0.19 and 0.57, respectively.

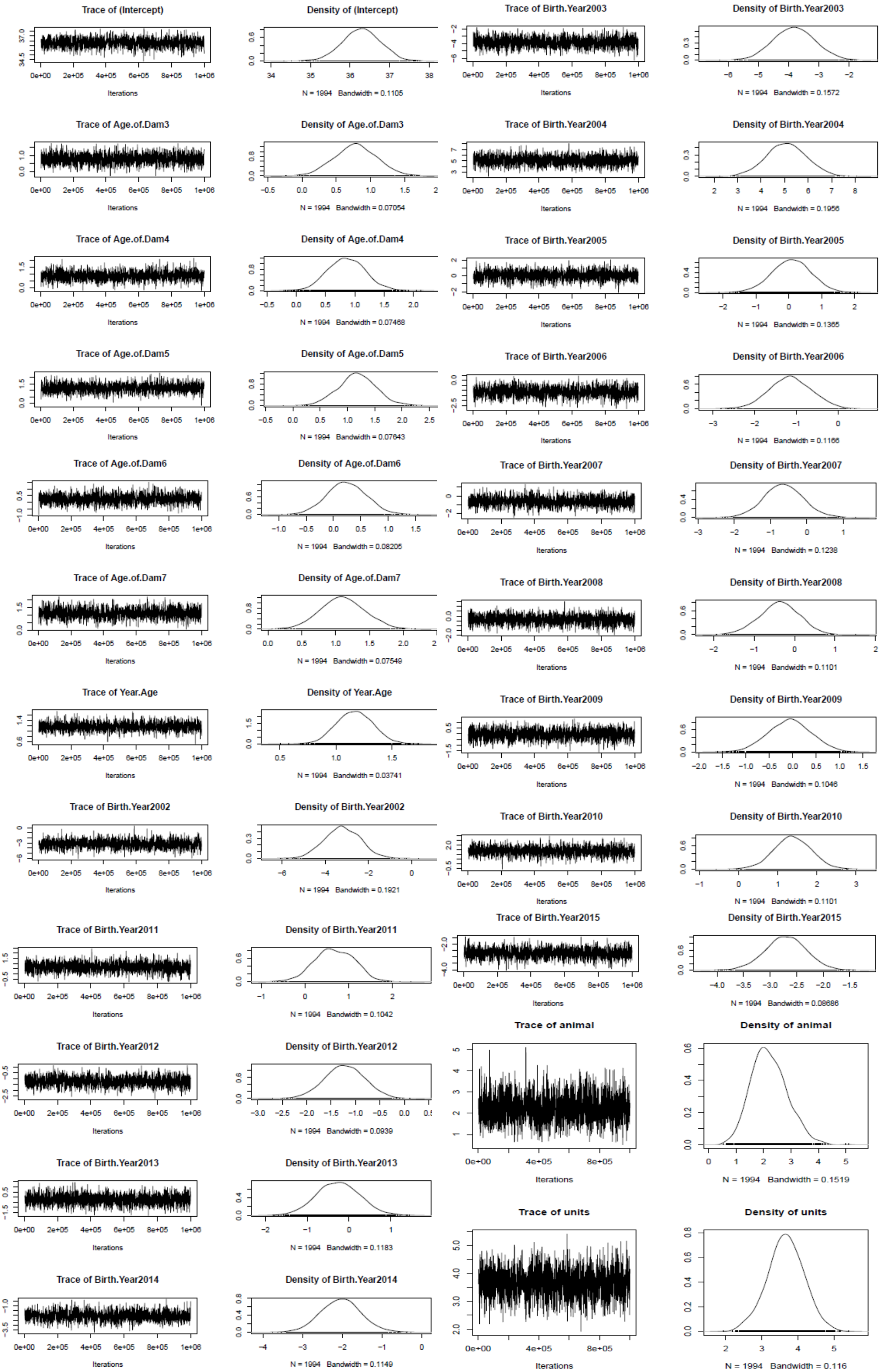

Figure A8.1 Chain trace plots (on the left) and approximated posterior density functions (on the right), obtained by the MCMC algorithm, for the intercept, age of dam, birth year, additive genetic (animal) and residual (units) in the animal model with yearling scrotal circumference as dependent variable