AGRICULTURAL UNIVERSITY OF ATHENS
FACULTY OF ANIMAL SCIENCE AND AQUACULTURE
DEPARTMENT OF ANIMAL BREEDING AND HUSBANDRY

Ph.D. THESIS

CONTEMPORARY METHODS OF ESTIMATION OF GENETIC PARAMETERS IN BROILER CHICKEN POPULATIONS

MANIATIS GERASIMOS

Consulting Committee:

Kominakis A. Assoc. Professor

Banos G. Professor

Demiris N. Lecturer

Athens, September 2012

ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΖΩΙΚΗΣ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΥΔΑΤΟΚΑΛΛΙΕΡΓΕΙΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΓΕΝΙΚΗΣ ΚΑΙ ΕΙΔΙΚΗΣ ΖΩΟΤΕΧΝΙΑΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΣΥΓΧΡΟΝΕΣ ΜΕΘΟΔΟΙ ΕΚΤΙΜΗΣΗΣ ΓΕΝΕΤΙΚΩΝ ΠΑΡΑΜΕΤΡΩΝ ΣΕ ΠΛΗΘΥΣΜΟΥΣ ΚΡΕΟΠΑΡΑΓΩΓΩΝ ΟΡΝΙΘΙΩΝ

ΜΑΝΙΑΤΗΣ ΓΕΡΑΣΙΜΟΣ

Εξεταστική Επιτροπή:

Άμπας Ζαφείρης Αν.Καθηγητής

Δεμίρης Ν. Λέκτορας

Ζέρβας Γ. Καθηγητής

Κομινάκης Α. Αν. Καθηγητής

Μπάνος Γ. Καθηγητής

Πολίτης Ι. Καθηγητής

Σκαράκης Γ. Καθηγητής

Αθήνα, Σεπτέμβριος 2012

CONTENTS

LIST OF TABLES AND FIGURES

ΣΥΓΧΡΟΝΕΣ ΜΕΘΟΔΟΙ ΕΚΤΙΜΗΣΗΣ ΓΕΝΕΤΙΚΩΝ ΠΑΡΑΜΕΤΡΩΝ ΣΕ ΠΛΗΘΥΣΜΟΥΣ ΚΡΕΟΠΑΡΑΓΩΓΩΝ ΟΡΝΙΘΙΩΝ

ΜΑΝΙΑΤΗΣ ΓΕΡΑΣΙΜΟΣ

Τμήμα Επιστήμης Ζωικής Παραγωγής και Υδατοκαλλιεργειών, Εργαστήριο Γενικής και Ειδικής Ζωοτεχνίας, Ιερά Οδός 255, ΤΚ 11855, email: acom@aua.gr

Περίληψη

Πραγματικά και προσομοιωμένα δεδομένα που ακολουθούσαν κανονική και διωνυμική κατανομή αναλύθηκαν χρησιμοποιώντας τη μέθοδο της υπό περιορισμό μέγιστης πιθανοφάνειας καθώς και δύο μεθόδους Bayes (MCMC και INLA). Τα αποτελέσματα της εργασίας δηλώνουν την χρησιμότητα της εφαρμογής των MCMC στη διερεύνηση συνδιακυμάνσεων μεταξύ τυχαίων επιδράσεων. Η μέθοδος INLA φαίνεται να αποτελεί μία εναλλακτική μέθοδο Bayes γρήγορη και χρήσιμη για ανάλυση δεδομένων που ακολουθούν κανονική κατανομή. Για την επιλογή του καταλληλότερου στατιστικού προτύπου ανάλυσης μία σειρά κριτηρίων έχει προταθεί. Εκτός από τα κριτήρια Akaike και BIC, στη συγκεκρίμενη μελέτη εφαρμόστηκε και το cAIC το οποίο λαμβάνει πιο αποτελεσματικά υπόψιν τους δραστικούς βαθμούς ελευθερίας και όπως προέκυψε από τις αναλύσεις είναι δυνατό να προτείνει διαφορετικά πρότυπο. Εξετάσθηκε η συνεισφορά των γενετικών επιδράσεων στο φαινόμενο του φυλετικού διμορφισμού του βάρους κρεοπαραγών ορνιθίων, χρησιμοποιώντας διμεταβλητά πρότυπα ανάλυσης αλλά και δευτερογενείς ιδιότητες που λάμβαναν υπόψη τη διαφορά και το λόγο των σωματικών βαρών αρσενικών και θηλυκών ατόμων. Για το σκοπό αυτό, χρησιμοποιήθηκε ένας μεγάλος αριθμός δεδομένων, αποτελούμενος από 203,323 ατομικές αποδόσεις σωματικών βαρών κρεοπαραγωγών ορνιθίων. Το καλύτερο πρότυπο περιελάμβανε προσθετικές γενετικές, μητρικές γενετικές και μητρικές περιβαλλοντικές επιδράσεις καθώς και τη συνδιακύμανση μεταξύ προσθετικών και μητρικών γενετικών επιδράσεων. Ο συντελεστής κληρονομικότητας δεν παρουσίασε διαφορές μεταξύ αρσενικών και θηλυκών (0.28 vs. 0.29). Μόνο ο συντελεστής συσχέτισης μεταξύ προσθετικών και μητρικών γενετικών επιδράσεων διέφερε ανάμεσα στα φύλα (-0.72 vs. -0.56 για αρσενικά και θηλυκά βάρη, αντίστοιχα), υποδεικύοντας έναν πιο έντονο ανταγωνισμό μεταξύ προσθετικών και μητρικών γενετικών επιδράσεων για τα αρσενικά. Τέλος, χρησιμοποιήθηκαν επιπλέον 35,595 αποδόσεις βαρών κρεοπαραγωγών ορνιθίων στις 35 ημέρες υπό δυσμενείς μικροβιακά συνθήκες για τη διερεύνηση της ύπαρξης αλληλεπίδρασης γονοτύπου περιβάλλοντος. Τα αποτελέσματα της μελέτης έδειξαν την ύπαρξη χαμηλού συντελεστή προσθετικής γενετικής συσχέτισης μεταξύ των δύο περιβαλλόντων, κυμαινόμενου μεταξύ 0,28 και 0,45 ανάλογα το πρότυπο, υποδηλώνοντας την ύπαρξη ισχυρής αλληλεπίδρασης γονοτύπου περιβάλλοντος για τη συγκεκριμένη ιδιότητα στον υπό εξέταση πληθυσμό.

1. General Introduction

As poultry consumption is projected to grow at 2.5 percent *per annum* to 2030 (FAO, 2008) there is an increasing demand for poultry meat and eggs in many parts of the world. This demand favors the industrialization of the production systems. The poultry sector is the most industrialized of all forms of livestock production and large-scale production is now widespread in many developing countries. Global production of chickens has experienced massive change and growth over the past 50 years. The commercial broiler and layer markets produce more than 40 billion birds annually to meet current worldwide consumer demands of more than 61 metric tons of meat and more than 55 million metric tons of eggs. Production has increased by 436% since 1970, more than 2.3 times and 7.5 times the corresponding growth in swine and beef, respectively (FAO, 2008).

Only a few multinational breeding companies currently remain as genetic suppliers of the majority of the commercial birds. While in the 1960s there were hundreds of breeding companies with significant market influence, by 1980, 13 broiler breeding companies remained and by 2001 only four independent groups with significant world market shares had survived (Besbes *et al.*, 2008). During the same time span, genetic improvement shifted from selection applied available chicken breeds to introduction of line specialization and crossbreeding, both stimulated by earlier developments in plant breeding. Literally, it was the degree of successful introduction of crossbreeding along with the ability to foresee the forthcoming changes in the market place that caused the number of breeding companies to reduce rapidly (Muir *et al.*, 2008).

Breeding efforts have dramatically increased meat poultry efficiency and it has been estimated that 80-90% of the progress over time has been made possible by improvement of the genetic potential (Havenstein *et al.*, 1994). The growth rate of modern broilers has roughly quadrupled since commercial breeding commenced in the 20th century. By the year 2000, the broilers' body weight reached 2.5 kg at 42 days of age with 1.75 feed conversion ratio (Besbes *et al.*, 2008). The body composition of the birds has also changed dramatically with broilers having more breast and thigh muscle, a relatively high proportion of abdominal fat and smaller spleen and heart weights than traditional lines (Sandercock *et al.*, 2009). A variety of selection traits are currently being used, such as hatching egg production, growth rate, feed efficiency, meat yield traits, liveability and leg and skeletal length (Besbes *et al.*, 2008). Differential selection pressure to the previous traits is being applied according to the market demands (live broilers, live/dress broiler or deboned market).

For broilers, further genetic gains are still pursued via multi-way approaches among which is the application of more sophisticated statistical methods as well as the application of genomic selection (GS). During the last few years, GS has greatly evolved and expectations are high. Although still costly, major breeding companies will use it to improve traits showing low levels of additive genetic variance such as disease resistance and robustness. At the same time, Quantitative Genetics is enriching its armamentarium with more sophisticated mathematical models as well as appropriate statistical procedures for estimation and model evaluation.

From the early developments such as the Henderson's mixed model equations (1959), Animal Breeding has traditionally been at the forefront of the application of sophisticated statistical methods. The lack of appealing of these methods along with the shift in funding away from quantitative genetics to almost exclusively molecular

genetics have led to a significant shortage of quantitative geneticists and another endangered species (Misztal, 2007) in need to be preserved.

Accurate genetic evaluation remains a topic of major interest, both scientifically and commercially. Specifically, accurate prediction of breeding values (BVs) is of great importance since a few selected animals have a major influence on the genetic progress of the whole population. A two-stage procedure is typically being employed: first, estimation of the variance components is performed followed by prediction of BVs. There are two ways to accomplish the above tasks. The first is based on the concept of the likelihood while the second on probability (Bayesian approach). For the frequentist breeder, the standard methods for VCE and prediction of BVs are REML (Patterson and Thompson, 1971) and BLUP (Henderson, 1975), respectively. Markov Chain Monte Carlo (MCMC) methods (Gelfand and Smith, 1990) have enormously increased the applicability of the Bayesian approach. These methods are computationally demanding and time consuming procedures and this has been the main hindrance to their application in animal breeding. Recently, a non-sampling based alternative to MCMC, the Integrated Nested Laplace Approximation (INLA) has been introduced (Rue *et al.*, 2009). Because INLA is based on direct numerical integration instead of simulations, it is much faster than MCMC and might represent an attractive alternative to REML for the animal breeder. Studies examining the relative merits of all the three methods (REML, Gibbs sampling and INLA) for VCE and estimation of BVs using field and simulated data following both normal and binomial distributions have been lacking from the literature. The second chapter was designed to provide valuable knowledge in this direction.

Various model selection procedures have been proposed thus far. Perhaps the most commonly used are the likelihood ratio test (LRTs) and two information criteria: the Akaike Information Criterion (AIC; Akaike, 1973) and the Bayesian Information Criterion (BIC; Schwarz, 1978). For the information criteria, a major issue when comparing different models is the appropriate choice of the penalty term when penalizing for the complexity of the models. The determination of the number of the model parameters is non-trivial when random effects are estimated using methods such as BLUP. On this occasion, model parameters could range from a small number of variance components to the complete number of random effects involved. Vaida and Blanchard (2005) defined a model evaluation criterion, the conditional AIC (cAIC) that effectively accounts for the complexity of the models by utilizing the effective degrees of freedom ρ (Hodges and Sargent, 2001). The use of cAIC as a model evaluation criterion has not hitherto been explored in the context of animal breeding. It is first introduced and studied in the third chapter of the present PhD thesis.

The accurate estimation of genetic parameters presumes sufficiently large data sets, with desirable data structure(s). In meat poultry, an exhaustive description of the various random effects that may affect phenotypic expression of traits such as body weight (BW) is only possible when there are enough half and full-sib families and adequate dam-offspring pairs with records. Only structures of this kind may allow for accurate and reliable estimation of the maternal genetic effects and the direct-maternal covariance. Very few attempts have been made to partition the maternal variance into genetic and environmental components, in poultry (Koerhuis and Thompson, 1997). BW at slaughter remains a trait of major economic importance since it relates to growth rate and meat yield. The trait in question shows considerable inter-sexual variation i.e. sexual dimorphism (SD) the genetic basis of which is rather poorly investigated. Chapters 3 and 4 were designed to provide reliable estimates of maternal (genetic and

environmental) effects, as well as of the direct-maternal covariance (correlation). Age (chapter 3) and sex-specific (chapter 4) analyses for the trait itself as well as for sex combinations that describe SD were estimated (chapter 4).

Most of the improvement process is carried out in breeding companies located in temperate regions. Products are then marketed worldwide – including tropical, semi-arid and arid regions where conditions are challenging in terms of climate, husbandry, feeds and feeding practices as well as hygiene conditions. A problem usually faced here is genotype by environment interaction (GEI). When present, the relative advantages of genotypes may differ from one environment to the other, in such a way that it could become necessary to choose specific genotypes for specific environments. In order to account for GEI, breeding companies either test their animals across a range of environments or establish satellite breeding programmes in various locations. GEI may arise from various environmental conditions such as nutritional (Havenstein *et al.*, 1994), climatic (e.g. Cahaner and Leenstra, 1992) or hygiene (e.g. Banos *et al.*, 2006) levels. The latter can potentially have a dramatic impact on the performance of broilers and there seems to be minimal relevant information reported in the literature, at least from a quantitative genetics viewpoint. The fifth chapter investigates the importance of GEI arising from hygiene conditions and examines whether sex-specific breeding policies are necessary to accommodate this interaction.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, Edition Petrov and Csaki, Budapest, Hungary.

Banos, G., Avendano, S. and Olori, V. 2006. Time dependent genetic parameters for broiler chicken body weight measured in selection and commercial environments. *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*, *Belo Horizonte, Brazil.*

Besbes, B., Tixier-Boichard, M., Hoffmann, I. and Jain, G.L. 2008. Future trends for Poultry Genetic resources. *International Poultry Conference*, Bangkok, November 2007. FAO.

Cahaner, A. and Leenstra, F.R. 1992. Effects of high temperature on growth and efficiency of male and female broilers from lines selected for weight gain, favorable feed conversion and high or low fat content. *Poultry Science* 71: 1237-1250.

Gelfand, A.E., and Smith, A.F.M. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85: 398–409.

Food and Agriculture Organization of the United Nations, 2008. *FAOStat*. Available at http://faostat.fao.org. Accessed June 1, 2008.

Geman, S., and Geman, D. 1984. Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–741.

Havenstein, G.B., Ferket, P.R., Scheideler, S.E. and Larson, B.T. 1994. Growth, liveability and feed conversion of 1957 vs. 1991 broilers when fed 'typical' 1957 and 1991 broiler diets. *Poultry Science* 73: 1785–1794.

Henderson, C.R., Kempthorne, O., Searle, S.R. and von Krosigk, C.M. 1959. The estimation of environmental and genetic trends from records subject to culling. *Biometrics* 15: 192-217.

Henderson, C.R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31: 423-447.

Hodges, J.S. and Sargent, D.J. 2001. Counting degrees of freedom in hierarchical and other richly-parameterized models. *Biometrika* 88: 367-379.

Koerhuis, A.N.M. and Thompson, R. 1997. Models to estimate maternal effects for juvenile body weight in broiler chickens. *Genetics Selection Evolution* 29: 225–249.

Misztal, I. 2007. Shortage of quantitative geneticists in animal breeding. *Journal of Animal Breeding Genetics* 124: 255–256.

Patterson, H.D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545-554.

Rue, H., Martino, S. and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B* 71(2):319-392.

Sandercock, D.A., Barker, Z.E., Mitchell, M.A. and Hocking, P.M. 2009. Changes in muscle cell cation regulation and meat quality traits are associated with genetic selection for high body weight and meat yield in broiler chickens. *Genetics Selection Evolution* 41: 8 doi:10.1186/1297-9686-41-8.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464

Sorensen, D. A., C. S. Wang, J. Jensen, and Gianola, D. 1994. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genetics Selection Evolution* 26: 333–360.

Vaida, F. and Blanchard, S. 2005. Conditional Akaike information for mixed-effects models. *Biometrika* 92: 351-370.

2. An old question still seeking for an answer: REML or Bayes?

2.1 Summary

REML has become the standard method of variance components estimation in animal breeding. Inference in Bayesian animal models is typically based upon Markov Chain Monte Carlo (MCMC) methods which are generally flexible but time consuming. Recently, a new Bayesian computational method, Integrated Nested Laplace Approximation (INLA), has been introduced for making fast non-sampling-based Bayesian inference for hierarchical Gaussian Markov models. This paper is concerned with the application and comparison of estimates provided by the three methods (REML, MCMC and INLA) using three representative programs (ASREML, WinBUGS and the R-package AnimalINLA) to the methods. A dataset comprising of 2,319 body weight records of a commercial line of broiler chicken was used. Both, a normally distributed trait i.e. body weight at 35 days of age (BW) and a binary response trait (after transformation of the Gaussian trait, assuming a 20% value as the threshold value), were explored. All model evaluation criteria suggested a purely additive animal model, in which the heritability estimates ranged from 0.31 to 0.34, for the Gaussian trait, and from 0.19 to 0.36 for the binary trait, depending on the software. WinBUGS revealed the existence of a small negative correlation between the additive genetic and maternal environmental effects (-0.20) that is usually neglected during REML-based analyses and may have an effect on parameter estimation. A simulation study was conducted based upon the results of the real dataset (Gaussian case), exploiting two scenarios of correlation between direct genetic and maternal environmental effects. Results suggest that while WinBUGS appeared to successfully accommodate more complicated structures between the various random effects, REML remains a fast and accurate procedure with general applicability. Although in need for further development, AnimalINLA seems a fast program for Bayesian Animal modelling, particularly suitable for inference of Gaussian traits.

2.2 Introduction

The Restricted Maximum Likelihood (REML) method (Patterson and Thompson, 1971) for unbalanced mixed models has been extensively used in animal breeding and has become the standard method for the estimation of variance components estimation. The Bayesian Markov chain Monte Carlo (MCMC) methods were introduced in quantitative genetics in the early 1990s (Wang *et al.* 1993; Sorensen *et al.* 1994), facilitated by the development of the Gibbs sampling procedure (Geman and Geman 1984; Gelfand and Smith 1990). The Gibbs sampler successively samples from conditional distributions of all parameters in a model in order to generate a random sample of the marginal posterior distribution, which is the target for Bayesian inference. MCMC methods represent the standard inference procedure for Bayesian animal models (Sorensen and Gianola, 2002) and through the years they have become an attractive alternative to REML. Recently, a non-sampling based alternative to MCMC, the Integrated Nested Laplace Approximations (INLA) has been introduced (Rue *et al.*, 2009). Using INLA, marginal posteriors for all parameters and random effects can be calculated. Because INLA is based on direct numerical integration instead of simulations, it is much faster than MCMC (Rue *et al.*, 2009). Furthermore, Holand *et al.* (2011) have developed an R package (AnimalINLA) making Bayesian animal models more accessible to animal breeders.

Several programs are available for MCMC methods, but very few provide a flexible environment. WinBUGS (Lunn *et al.* 2000) is the most well-developed and

general-purpose Bayesian software available to date. It has an interactive environment that enables the user to specify models that need to be compiled before starting the Gibbs sampling. Convergence diagnostics, model comparisons, e.g. via DIC, and other useful plots and diagnostics are available. Several distributions can be used for modelling the observations as well as priors, while full conditional distributions are automatically constructed and the appropriate MCMC algorithm for sampling is chosen (Lunn *et al.*, 2000). The importation of the animals' genetic relationships has been an issue in WinBUGS and several attempt to this direction have been made (Damgaard, 2007; Waldmann, 2009). However, these methods either required prior transformation of the data using complex code or did not provide a generic procedure independent of the data structure. To accommodate the problem, Gorjanc (2010) suggested the use of the inverse of the numerator relationship matrix A^{-1} directly through the diagonal values of W^{-1} matrix, where $A^{-1} = (T^{-1})'W^{-1}T^{-1}$ (Henderson, 1976; Quaas, 1976).

The primary goal of this study was to investigate the relative merits of the three methods (REML, Gibbs sampling and INLA) in the context of animal breeding, using representative programs such as ASREML 3.0 (Gilmour *et al.*, 2009), WinBUGS and AnimalINLA. For this purpose, a Gaussian and a binary trait were both explored and variance components as well as the genetic parameters along with breeding values across the three methods were estimated and compared. Moreover, a simulation study was conducted in the case of a Gaussian trait, under two different scenarios of correlation (r_{uc}) between the direct genetic and the c^2 effects.

2.3 Materials and methods

2.3.1 Data description

Data on body weight (BW) at 35 days of age from a broiler line were made available by Aviagen Ltd. The dataset consisted of 2,319 records comprising 1,171 males and 1,148 females in 41 hatch weeks, while the pedigree included a total of 2,456 animals. To make results directly comparable, all phenotypic values were standardized to the standard normal distribution via $y = \frac{y_0 - \bar{y}}{\sigma_{y_0}}$, where y ~ N(0, 1) the standardized BW, y_0 the original phenotypic values of BW, $\bar{y}$ the mean BW in the population and σ_{y_0} the standard deviation of BW. A preliminary analysis of variance showed that the statistically significant fixed effects ($P<0.05$) included hatch week and sex. Hence, these fixed effects were included in all models. Every female was mated with at least 2 males producing from 2 to 57 offspring with records, while the 32 sires were mated with 2 to 7 females and produced 2 to 97 offspring. This structure enables the estimation of c^2 effects through proper modeling and this was pursued in both the real and simulated data.

A binary response trait was also built, using the original BW values and a threshold at the highest 20% phenotypic values. Thus, the new variable y_{20} followed the Bernoulli distribution with values 0 and 1 denoting low and high weight, respectively. In this data set, only the gender of the animals was statistically significant ($P<0.05$) and was included in analyses as the only fixed effect.

2.3.2 Statistical analysis

2.3.2.1 Gaussian data trait

Four animal models were considered for BW. Model M_1 was a purely direct additive animal model, while model M_2 allowed for the inclusion of maternal environmental effects. Models M_3 and M_4 were as model M_2, but with zero and non-zero σ_{uc}, respectively. In summary, the models in matrix notation were as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \quad (M_1)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_c c} + \mathbf{e} \quad (M_2)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_c c} + \mathbf{e}, \text{ with } \operatorname{cov}(u, c) = 0 \quad (M_3)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_c c} + \mathbf{e}, \text{ with } \operatorname{cov}(u, c) = \sigma_{uc}\mathbf{I} \quad (M_4),$$

where $\mathbf{y}$ = n×1 vector of observations (n = number of records = 2,319), $\mathbf{b}$ = p×1 vector of fixed effects (p = number of fixed effects classes = 42), $\mathbf{u}$ = q×1 vector of direct additive genetic effects (q = number of additive effects = 2,456), $\mathbf{c}$ = k×1 vector of maternal environmental effects (k = number of dams with offspring= 105), $\mathbf{e}$ = n×1 vector of residuals; $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{Z_c}$ denote the incidence matrices relating the observations to the corresponding fixed and random effects; and $\mathbf{A}$ the additive relationship matrix. The vector of direct genetic effects was assumed to follow the normal distribution: $\mathbf{u} \sim N(\mathbf{0}_n, \sigma_u^2 \mathbf{A})$, where $\mathbf{0}_n$ denotes a n×1 vector of 0s and σ_u^2 denotes to the direct genetic variance. The maternal environmental effects were assumed to follow a normal distribution given by: $\mathbf{c} \sim N(\mathbf{0}_k, \sigma_c^2 \mathbf{I}_k)$, where $\mathbf{I_k}$ is an identity matrix of order k and σ_c^2 the maternal environmental variance. Finally, residuals for the two traits were assumed normal as follows: $\mathbf{e} \sim N(\mathbf{0}_n, \sigma_e^2 \mathbf{I}_n)$, where σ_e^2 is the residual variance.

From a Bayesian perspective, the data $\mathbf{y}$ are assumed to be a realization from $\mathbf{y} \mid \mathbf{b}, \mathbf{u}, \sigma_e^2 \sim N(\mathbf{Xb} + \mathbf{Zu}, \sigma_e^2 \mathbf{I}_n)$ and $\mathbf{y} \mid \mathbf{b}, \mathbf{u}, \mathbf{c}, \sigma_e^2 \sim N(\mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_c c}, \sigma_e^2 \mathbf{I}_n)$ for models M_1 and M_2, respectively, where the unobserved vector (q×1) of the additive genetic effects is assumed to follow the multivariate normal distribution: $\mathbf{u} \mid \mathbf{A}, \sigma_u^2 \sim N(0, \sigma_u^2 \mathbf{A})$ and the vector of the maternal environmental $\mathbf{c}$ (k×1) the multivariate normal $\mathbf{c} \mid \sigma_c^2 \sim N(0, \sigma_c^2 \mathbf{I}_k)$. The vector of fixed effects $\mathbf{b}$ (p×1) was partitioned into two sub-vectors, denoting hatch (h) and sex (s). It was assumed that both sub-vectors followed univariate normal according to: $\mathbf{h} \mid \sigma_h^2 \sim N(0, \sigma_h^2)$ and $\mathbf{s} \mid \sigma_s^2 \sim N(0, \sigma_s^2)$. Two additional Bayesian animal models were explored that included a covariance between additive genetic and maternal environmental effects. The vector of the data $\mathbf{y}$ for both models was assumed to be:

$\mathbf{y} \mid \mathbf{b}, \mathbf{u}, \mathbf{c}, r, \sigma_e^2 \sim N(\mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_c c}, \sigma_e^2 \mathbf{I}_n)$, where the correlation was $r = \dfrac{\operatorname{cov}(u, c)}{\sigma_u \sigma_c}$ and $r = 0$ for models M_4 and M_3, respectively.

Gelman (2006) investigated the statistical properties of different priors on variance components and found that a uniform prior on the standard deviation is a reasonable choice in a number of situations. Therefore, a vague uniform prior was utilised for the standard deviation of the additive genetic effects $\sigma_u \sim U(0, 100)$. The priors for c^2 and the residual variance σ_e^2 were Inverse-gamma (0.001, 0.001). A normal prior was

assumed for the fixed effects $\mathbf{b_i} \sim N(0, 10^4)$. Estimates of heritability (h^2) as well as c^2 were calculated as ratios of the estimates of direct additive (σ^2_u) and maternal environmental (σ^2_c) variances to the phenotypic variance (σ^2_p), respectively. The phenotypic variance accounts for the sum of all variance components, according to the model.

2.3.2.2 Simulation study

A simulation study was also conducted with population structure emulating the pedigree and the variance components of the real data. In total, 20 sires and 70 dams were used in the pedigree and 2240 progeny with records were simulated. Each sire was assumed to mate to 7 females, while every dam produced offspring with 2 different males. Each full-sib family comprised of 16 offspring. Direct genetic effect for founder i (1,…,90) was drawn as $\mathbf{u}_i \sim N(0,\sigma^2_u)$, while the maternal environmental effect of dam j (1,…,70) was $\mathbf{c}_j \sim N(0,\sigma^2_c)$, with $\sigma^2_u = 7$ and $\sigma^2_c = 3$. Two scenarios were explored regarding the correlation between the direct genetic and the c^2 effects (r_{uc}): a) r_{uc} = -0.2 (low) and b) r_{uc} = -0.8 (high). The direct genetic effects of offspring i (1, …,2240) were calculated by: $u_i = \frac{1}{2}(u_j + u_k) + ms_i$, where u_j and u_k denote direct genetic effects of dam and sire, respectively, while ms_i represented the Mendelian sampling deviation drawn conditional upon the c^2 effects: $ms_i \mid c_i \sim N(\frac{\sqrt{0.5\sigma^2_u}}{\sigma_c} r\, c_i, (1-r^2)0.5\sigma^2_u)$. The total phenotypic variance was estimated according to: $\sigma^2_p = \sigma^2_u + \sigma^2_c + \sigma^2_e$. The residuals were sampled as $e_i \sim N(0,\sigma^2_e)$, where $\sigma^2_e = 32$, thus resulting in $\sigma^2_p = 42$, $h^2 = 0.17$ and $c^2 = 0.07$.

Note that all the above values for the various variance components were deliberately chosen to have: a) a marginal but statistically significant contribution of the c effects, and b) a negative r_{uc}. Thus the model of analysis of the simulated data would be M_4. A common mean was assumed to be the only fixed effect for the simulated trait. In total, 30 samples from each scenario were generated. These samples were then analyzed via models M_1-M_2 (ASREML and AnimalINLA) and M_2-M_4 (WinBUGS). The mean squared error (MSE) was employed to quantify the performance of the predictors throughout along with the coverage of interval estimates. The MSE was computed as follows: $MSE = \frac{\sum_{i=1}^{N} ((\hat{\theta} - \theta)^2 + \mathrm{var}(\hat{\theta}))}{N}$ where θ stands for the true, $\hat{\theta}$ for the estimated parameter, $\hat{\theta} - \theta$ corresponds to *bias* and N = 30 is the number of samples.

2.3.2.3 Binary data trait

Initially, a simple animal model was fitted via REML, considering y_{20} as a normally distributed trait. Subsequently, an animal threshold model (Wright, 1934; Dempster and Lerner, 1950) was used for the analysis of the binary variable. In this analysis, the observed binary variable y20 is related to an underlying unobservable continuous variable, called liability λ (Gianola, 1982; Gianola and Foulley, 1983), such that the observed binary responses (y_{20}) are the result of the following relationship:

$$y_{20i} = \begin{cases} 0 \; if \; \lambda_i \leq \tau \\ 1 \; if \; \lambda_i > \tau \end{cases},$$

where τ is a fixed threshold, and y_{20i} corresponds to observation i. Several link functions (logit, probit, cloglog) can be applied to link the binary variable to the underlying scale (Gilmour *et al.*, 2009). In our study, the logit link function was used: $\lambda = \log(\frac{\mu}{1-\mu})$, where μ is the probability of success and λ the vector of linear predictors of the liability on the underlying scale. An animal model was assumed for λ such as: $\lambda = Xb + Zu + e$. A uniform prior was assumed here for the standard deviation of the additive genetic effects on the underlying scale $\sigma_u \sim U(0, 100)$. In the threshold models, the residuals have fixed variance with values depending on the chosen link function: $\sigma_e^2 = 1$ and $\sigma_e^2 = \frac{\pi^2}{3} \approx 3.29$ for probit and logit, respectively (Gilmour *et al.*, 2009). On the logit scale, the heritability h^2 was estimated by: $h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\pi^2}{3}}$.

2.3.2.4 Model evaluation criteria

According to the method applied, model comparison was based on four evaluation criteria: the Akaike Information Criterion (AIC; Akaike, 1973), the Bayesian Information Criterion (BIC; Schwarz, 1978), the conditional Akaike Information Criterion (cAIC; Vaida and Blanchard, 2005) and the Deviance Information Criterion (DIC; Spiegelhalter *et al.*, 2002). All criteria are based upon the computation of the deviance (D): $D = -2\log(p(y|\hat{\theta})) = -2\log L$, where θ denotes the *px1* vector of the model parameters and $p(y|\hat{\theta})$ the likelihood of the data y evaluated at the maximum likelihood estimate $\hat{\theta}$. While LRTs suggest the direct comparison of LogLs between the various nested models, AIC, BIC and cAIC suggest penalizing the deviance by appropriate complexity terms. According to Akaike (1973) the appropriate term for penalizing the deviance is twice the number of the model parameters p *i.e.* $AIC = -2\log L_i + 2p$ while Schwarz (1978) suggested the proper penalization term to be $p\log(n)$: $BIC = -2\log L_i + p\log n$, where n is the number of observations. However, the determination of the number of the model parameters is non-trivial when random effects are of interest and are being estimated using methods such as BLUP. For such cases the AIC is shown (Crainiceanu and Ruppert, 2004) to be asymptotically biased. In addition, Greven and Kneib (2010) showed that in linear mixed models AIC is a biased estimator of the Akaike information due to the non-open parameter space and the lack of independence between observations. An asymptotically unbiased criterion is the cAIC defined by Vaida and Blanchard (2005) as $cAIC = -2\log L_i + 2\rho$, where ρ are the effective degrees of freedom (Hodges and Sargent, 2001), given by the trace of the hat-matrix **H**. A Bayesian analogue of ρ, the the effective number of parameters, p_D, has been proposed earlier and is used in the DIC (Spiegelhalter *et al.*, 2002): DIC = D + $2p_D$. In all criteria, models with smallest values are to be preferred denoting better fit to the data.

2.4 Results

2.4.1 Gaussian data trait

Table 2.1 summarizes the estimated variance components and genetic parameters of BW, along with likelihoods, ρ and model evaluation criteria. Heritability for BW ranged from 0.15 to 0.34, while c^2 accounted for 0-0.08 of the total phenotypic variance, depending on the model and the method applied. All evaluation criteria, regardless of the method considered, concur in the choice of a purely additive animal model without the inclusion of the c^2 effects. During M_1, heritability estimates ranged slightly among the methods from 0.31 (ASREML) to 0.34 (AnimalINLA), while confidence and credibility intervals between ASREML and the Bayesian programs always coincided.

Under model M_2, REML-based estimates were significantly different than those obtained from the two Bayesian approaches. In this case, REML direct heritability was seriously underestimated (0.15) when contrasted to MCMC and INLA methods (0.31 and 0.32, respectively). Furthermore, while c^2 was 0.07 ($\pm$ 0.03) in REML, no detectable variance due to c^2 was estimated during the Bayesian methods. As a result, the sum of the additive and the c^2 effects given as a proportion of the phenotypic variance was significantly lower during REML (0.22) when compared to Bayesian methods (0.31-0.32). Such a paradox may arise from covariances between the various random effects. To test for such a hypothesis we fitted model M_4 that accounted for a covariance between the additive genetic and the maternal environmental effects. This could be effectively modeled only via the WinBUGS software. Under model M_4 (Table 2.1), h^2 and c^2 estimates were comparable (0.17 and 0.08, respectively) to ASREML estimates (for model M_2), while the covariance in question was not statistically significant (0.04±0.07). A negative additive genetic-maternal environmental covariance was detected (-0.20), although with large standard error (0.30) that did not allow for firm conclusions. When this covariance was forced to 0 (model M_3) as in ASREML, estimates (0.28 and 0.02 for h^2 and c^2, respectively) approached those obtained (for model M_2) by WinBUGS and AnimalINLA.

To further quantify implications of model/method evaluation on selection decisions, Pearson as well as rank correlations of animals' EBVs and the percentage of common animals selected were calculated for all models and methods applied (results not shown). With regard to the EBVs, Pearson (and rank) correlations were extremely high (0.97, 0.99) when the focus was on the EBVs of all animals and/or a proportion of the best 20%. During this phase, an additional advantage of the WinBUGS software is its ability to estimate (via the Rank tool) the uncertainty associated with the ranking of the individuals from the posterior distributions of the EBVs. Figure 1 presents twelve selected examples from the posterior distribution of the EBV ranks, with four animals from the top, middle and low end of the spectrum, respectively. These ranks are based upon the whole posterior density and properly account for characteristics like the variance and skewness of the posterior. Both, a 95% rank interval as well as the median rank are provided, thus presenting an easy and flexible way of animal selection. We illustrate, among other aspects, the large uncertainty associated with selecting among similar animals. In order to account for uncertainty in the (AS)REML context, ranking of animals was derived by using the z-scores: $z_i = \frac{u_i}{se_i}$, where u_i and se_i the EBV and standard error of EBV of the animal i, respectively. Also here, rank correlations were remarkably high ranging from 0.96 to 0.99 among all methods and models considered. Furthermore, standard errors of the EBVs and solutions for the fixed effects were

comparable among the methods with no statistically significant differences. In general, all models and methods suggested the same animals (96-99% common amongst).

2.4.2 Simulation study

Descriptive statistics of the simulated data and the estimators across models and methods are in Table 2.2. Average values of the simulated data were equal to the true with h^2 and c^2 equal to 0.17 and 0.07, respectively. Note that during simulations c^2 was statistically significant. Using model M_1 under either ASREML or AnimalINLA resulted always in inflated predictions for the true parameters. While the true heritability was 0.17, the estimated heritability ranged from 0.35 to 0.51 with a tendency for inflated estimates particularly in AnimalINLA and under the strongly negative r_{uc} scenario for both software packages (ASREML and AnimalINLA). Overestimation of the direct heritability was due to both higher estimates of the additive genetic variance and lower estimates of the residual variances.

Estimates under model M_2 were in close proximity to the true values only in the case of ASREML and the low r_{uc} scenario (h^2=0.15, c^2=0.07). Slightly higher estimates for h^2 and c^2 were observed in ASREML in the extreme r_{uc} case (h^2=0.21, c^2=0.08). Under AnimalINLA, the respective h^2 estimator was seriously inflated (h^2=0.34) due to overestimation of the additive genetic effects and failure to account for the c^2 effects. This trend was more evident in the strong vs. the low r_{uc} scenario. WinBUGS estimates for Model M_2 under the high r_{uc} scenario were slightly better than those obtained by AnimalINLA. Finally, model M_4 was fitted via WinBUGS (Table 2.2) for the extreme scenario (r_{uc}=-0.8). Although a statistically significant r_{uc} was detected here (as high as -0.60), h^2 and c^2 were systematically overestimated.

The MSEs across models and methods are presented in Table 2.3. Irrespectively of the method and/or model, MSEs were significantly lower in low vs. the high correlation scenario. Furthermore, better estimates as reflected in terms of lowest MSEs were attained in ASREML using M_2 model under the low correlation. The MSE of model M_2 in ASREML was the lowest, whereas the MSE of model M_1 in AnimalINLA was the highest. Interestingly, even under the strongly negative r_{uc} scenario, using model M_2 in ASREML resulted in lower MSE, indicating that ignoring r_{uc} during modelling can result in effective estimators for the variance components under this particular software. The WinBUGS software, although able to account for the specific correlation, exhibited the highest MSE of σ_e^2, with analogous effect on the estimators of h^2 and c^2. All other parameters (σ_u^2 and σ_c^2) estimated via model M_4 in WinBUGS had relatively low MSE.

The coverage of interval estimates for the three models and the respective methods of analysis are shown in Table 2.4. To construct Bayesian 95% intervals, the quantiles of the relevant posterior distributions (as estimated by MCMC and INLA) were used. ASREML's intervals were constructed based on asymptotic normality of the maximum likelihood using $\hat{\theta} \pm 1.96 * se(\hat{\theta})$, where *se* denotes the estimated standard error of the parameter. In the case of low r_{uc} the best coverages were given by ASREML and model M_2, although intervals were narrower than the Bayesian methods. In contrast, WinBUGS exhibited the best coverage performance in the case of the high r_{uc}, in expense of wider intervals. AnimalINLA experienced difficulty in attaining nominal

coverage of interval estimates when model M_1 was assumed as well as under the scenario of the strongly negative r_{uc}. Furthermore, AIC suggested model M_2 analysed via ASREML in 60% of the samples under the low correlation scenario and in 33.33% under the high one. In contrast, DIC via AnimalINLA suggested under both scenarios less often (26.67% and 40.00% for low and high scenario, respectively) model M_2, which included the c^2 effects. Finally, DIC via WinBUGS favored the true model that incorporated the r_{uc} in 76.67% of the samples.

2.4.3 Binary data trait

The estimated variance components and genetic parameters of y_{20} for a purely additive animal model across the three methods are presented in Table 2.5. A model incorporating c^2 effects was also fitted, however convergence was not achieved under any method applied. In (AS)REML, heritability on the observed scale (h_O^2) was estimated as high as 0.10, while the respective estimate on the underlying scale was significantly higher (h_U^2=0.19). Using the classical formula (Dempster and Lerner, 1950), the ratio between the two estimates would be: $\frac{h_o^2}{h_U^2} = \frac{[z(x_p)]^2}{p(1-p)}$, where p is the level of incidence and $z(x_p)$ is the ordinate of a standard normal curve cutting off an area equal to p. For p = 0.2 (as in here) the ratio is $\frac{h_o^2}{h_U^2} \approx 0.5$ in full agreement with our results. Estimates from AnimalINLA were comparable to those of ASREML ($h_U^2 = 0.21$). Interestingly, the WinBUGS heritability estimate was significantly higher (up to 0.36) approaching the original h^2. Differences were also detected on the confidence/credibility intervals of the point estimates of the additive variance as well as the heritability on the underlying scale. More specifically, the credibility interval for h_U^2 given by WinBUGS was in the region of (0.21, 0.56), while that of AnimalINLA was in (0.13, 0.30) and more profoundly the confidence interval given by ASREML was (0.09, 0.29).

Rank correlations were also estimated here for the three methods (results not shown?). Although statistically significant differences were detected between the estimates, obtained by WinBUGS and the other two programs, rank correlations remained high ranging from 0.92 to 0.99. In addition, the proportion of common animals selected among the three methods exceeded 93% suggesting minor implications of method usage on selection decisions.

2.5 Discussion

The theoretical aspects and advantages of REML and MCMC methods for fitting hierarchical multilevel models, such as the animal model, have been extensively explored in literature (see Browne and Draper, 2006 for a thorough simulation study). However, this is the first study attempting to compare the efficiency and accuracy of three methods of variance components estimation and breeding values within the context of animal breeding. Our main concern was the practical aspects of applicability of three available typical software programs for the standard animal breeder.

From a frequentist's point of view, the standard method entails the use of the REML and BLUP methods. In the present study, the ASREML (Gilmour *et al.*, 2009)

software was employed. The software is free for academic purposes (at least the Windows-based; W-ASREML), is stable, fast, can handle many different models, data structures and thousands of data records. In addition, the necessary files are not very complicated to construct while a valuable manual, containing a lot of information and numerous examples, is available for the animal breeder. For binary trait modelling, a variety of link functions (logit, probit, cloglog) can be chosen. An obvious obstacle when using commercialized programs deals with their limited flexibility i.e. the inability to model complex structures between (random) effects. A good example here was the existence of a negative correlation between u and c that could not be appropriately accommodated within the context of a typical REML package. This covariance is usually treated as non existent (assumed of 0). Modelling of the covariance in question was only possible in WinBUGS, which enables the exploration of possible correlation structures between the various random effects, a potentially helpful feature in testing assumptions of the standard animal model. This program allows for the application of a large group of competing models and Bayesian model evaluation criteria (Sorensen and Gianola, 2002). A further important attribute of the WinBUGS program is the rank tool, which can simultaneously incorporate the uncertainty associated with the ranking of the individuals, thus assisting animal selection. Bayesian methods, such as the MCMC implemented in the WinBUGS, can be especially useful in complex situations at the cost of being computationally expensive and time consuming. For our data, after burn-in of 10,000 iterations a total number of 1,000,000 iterations and a thinning interval of 20 were applied. Based on graphical inspection of the trace and autocorrelation plots it was concluded that convergence was achieved, yielding a total sample of 50,000 iterations. This procedure took approximately 14 to 16 hours, depending on the model.

Present results based on the simulations have shown that: a) failure to account for statistically significant random effects apart from the direct genetic effects, results in inflated direct heritability estimates. This is the case because the additional random effects are accounted as direct genetic effects. This is a rather common finding across studies (e.g. Kushwaha *et al.*, 2009; Tosh *et al.*, 2010) and implies the need of using exhaustive models when analyzing data, b) even when an existent correlation between random effects is not taken into account during modelling, (AS)REML can provide effective point estimates of the variance components. Furthermore, WinBUGS flexibility and effectiveness in accounting for r_{uc} proved its usefulness for the Animal Breeder in unveiling such correlations of interest.

For the Bayesian animal breeder, AnimalINLA has proved a surprisingly time efficient experience. It took less than 10 sec to produce the required posterior distributions while providing comparable estimates with the other packages. Although computationally efficient, this R-package has revealed some disadvantages. Firstly, no more than 4,000 records could be incorporated into the animal model, probably due to compatibility problems with Windows. Secondly, as Holand *et al.* (2011) have also stated, AnimalINLA might give biased posteriors for the additive genetic variance in the case of a binary distributed trait. Finally, it is not as flexible in modelling as the WinBUGS, for the inexperienced in programming animal breeder and the manual is currently rather short.

In conclusion, WinBUGS can be of great assistance to the animal breeder, because of its flexibility to modelling complex models while unravelling existent data structures that the usual REML-based packages neglect. Within the animal breeding context, its applicability remains rather limited since only small to moderate data sets or

populations can be handled efficiently with respect to time. AnimalINLA appears a promising future perspective for the animal breeder dedicated to the Bayesian scholar since it is remarkably fast. It seems, however, to be a package still under development. Our own experience on large data sets have shown that ASREML can effectively handle analyses for up to 500,000 records and related pedigree structures, being stable, fast (<1 h) and mostly independent of initial values. Furthermore, as the simulation results have shown, even when an extreme covariance between random effects is neglected, it can provide estimates of the parameters in question with relatively small bias and error. Given all the above, the (AS)REML remains the gold standard for the heavy duty animal breeder.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, Edition Petrov and Csaki, Budapest, Hungary.

Browne, W. J. and Draper, D. 2006. A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis* 1: 473-514.

Crainiceanu, C. and Ruppert, D. 2004. Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society, Series B* 66: 165–85.

Daamgard, L.H. 2007. Technical note: How to use Winbugs to draw inferences in animal models. *Journal of Animal Science* 85: 1363-1368.

Dempster, E.R., and Lerner, I.M. 1950. Heritability of threshold characters. *Genetics* 35:212–235.

Gelfand, A. E., and Smith, A.F.M. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85: 398–409.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1: 515–534.

Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–741.

Gianola, D. 1982. Theory and analysis of threshold characters. *Journal of Animal Science* 54: 1079–1096.

Gianola, D. and Foulley, F. L. 1983. Sire evaluation for ordered categorical data with a threshold model. *Genetics Selection Evolution* 15: 201–224.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. 2009. ASReml User Guide Release 3.0. *VSN International Ltd*, Hemel Hempstead, UK.

Gorjanc, G. 2010. Flexible Bayesian Inference of Animal Model Parameters Using BUGS Program. Contribution for 9th WCGALP.

Greven, S. and Kneib, T. 2010. On the behaviour of marginal and conditional Akaike Information Criteria in Linear Mixed Models. *Biometrika* 97: 773-789.

Henderson, C. R. 1976. A simple method for the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32: 69–83.

Hodges, J.S. and Sargent, D.J. 2001. Counting degrees of freedom in hierarchical and other richly-parameterized models. *Biometrika* 88: 367-379.

Holand, A.M, Steinsland, I., Martino, S. and Jensen, H. 2011. Animal models and Integrated Nested Laplace Approximations (Submitted)

Kushwaha, B.P., Mandal, A., Arora, A.L., Kumar, R., Kumar, S. and Notter, D.R. 2009. Direct and maternal (co)variance components and heritability estimates for body weights in Chokla sheep. *Journal of Animal Breeding and Genetics* 126: 278-287.

Lund, M.S., Jensen, J. and Petersen, P. H. 1999. Estimation of genetic and phenotypic parameters for clinical mastitis, somatic cell deviance, and protein yield in dairy cattle using Gibbs sampling. *Journal of Dairy Science* 82: 1045–1051.

Lunn, D.J., Thomas, A., Best, N. and Spiegelhalter, D. 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 10: 325–337.

Mrode, R., and Thompson, R. 1989. An alternative algorithm for incorporating the relationships between animals in estimating variance components. *Journal of Animal Breeding Genetics* 106: 89–95.

Patterson, H. D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58: 545-554.

Quaas, R. L. 1989. Transformed mixed model equations: A recursive algorithm to eliminate A−1. *Journal of Dairy Science* 72: 1937–1941.

Rue, H., Martino, S. and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B* 71(2): 319-392.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464.

Sorensen, D.A., Wang, C.S., Jensen, J. and Gianola, D. 1994. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genetics Selection Evolution* 26: 333–360.

Sorensen, D.A., Andersen, S., Gianola, D. and Korsgaard, I. 1995. Bayesian inference in threshold models using Gibbs sampling. *Genetics Selection Evolution* 27: 229–249.

Sorensen, D., and Gianola, D. 2002. Likelihood, Bayesian and MCMC methods in quantitative genetics. Springer-Verlag, New York.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. 2002. Bayesian measures of model complexity and fit *Journal of the Royal Statistical Society, Series B* 75(4), 583-639 (with discussion).

Tosh, J.J., Garber, A.F., Trippel, E.A. and Robinson, J.A.B. 2010. Genetic, maternal, and environmental variance components for body weight and length of Atlantic cod at 2 points in life. *Journal of Animal Science* 88: 3513-3521.

Vaida, F. and Blanchard, S. 2005. Conditional Akaike information for mixed-effects models. *Biometrika*, 92, 351-370.

Waldmann, P. 2009. Easy and flexible bayesian inference of quantitative genetic parameters. *Evolution* 63: 1640-1643.

Wang, C.S., Rutledge, J. J. and Gianola, D. 1993. Marginal inference about variance components in a mixed linear model using Gibbs sampling. *Genetics Selection Evolution* 21: 41–62.

Wright, S. 1934. An analysis of variability in number of digits in an inbred strain of Guinea pigs. *Genetics* 19: 506–536.

Tables and Figures

Table 2.1 Estimates of variance components, genetic parameters, log-likelihoods and model evaluation criteria for the typical normally distributed body weight (BW) of broiler chicken at 35 days of age using three different methods software (ASREML, WinBUGS and AnimalINLA).

Software	Model		σ^2_u	σ^2_c	σ_{uc}	σ^2_e	σ^2_p	h^2	c^2	$\frac{\lvert\sigma_{uc}\rvert}{\sigma^2_p}$	r_{uc}	logL	AIC	BIC	cAIC/ DIC	ρ/ p_D
ASREML	M_1	Mean {CI}	0.133 {0.09, 0.18}	-	-	0.302 {0.27, 0.34}	0.434 {0.40, 0.47}	0.31 {0.21, 0.41}	-	-	-	-1,653	3,307	3,313	4,044	369
	M_2	Mean {CI}	0.065 {0.01, 0.13}	0.029 {0.01, 0.05}	-	0.335 {0.30, 0.38}	0.429 {0.40, 0.46}	0.15 {0.01, 0.29}	0.07 {0.01, 0.13}	-	-	-1,651	3,306	3,317	4,182	440
WinBUGS	M_1	Mean {CI}	0.139 {0.09, 0.20}	-	-	0.298 {0.26, 0.33}	0.437 {0.40, 0.48}	0.32 {0.22, 0.43}	-	-	-	-1,622	-	-	4,302	529
	M_2	Mean {CI}	0.134 {0.07, 0.20}	0.001 {0, 0.03}	-	0.300 {0.26, 0.34}	0.435 {0.38, 0.47}	0.31 {0.18, 0.43}	0 {0, 0.06}	-	-	-1,632	-	-	4,304	520
	M_3	Mean {CI}	0.126 {0.03, 0.20}	0.007 {0, 0.05}	-	0.305 {0.26, 0.36}	0.438 {0.41, 0.48}	0.28 {0.07, 0.43}	0.02 {0, 0.11}	-	-	-1,680	-	-	4,310	476
	M_4	Mean {CI}	0.069 {0.01, 0.15}	0.032 {0, 0.12}	-0.014 {-0.08, 0.01}	0.321 {0.20, 0.37}	0.410 {0.31, 0.47}	0.17 {0.04, 0.35}	0.08 {0, 0.36}	0.04 {0, 0.19}	-0.20 {-0.88, 0.44}	-1,795	-	-	4,305	358
AnimalINLA	M_1	Mean {CI}	0.152 {0.11, 0.21}	-	-	0.297 {0.26, 0.33}	0.449 {0.37, 0.54}	0.34 {0.23, 0.45}	-	-	-	-	-	-	4,289	-
	M_2	Mean {CI}	0.143 {0.10, 0.21}	0.004 {0, 0.03}	-	0.302 {0.26, 0.34}	0.449 {0.37, 0.57}	0.32 {0.23, 0.44}	0 {0, 0.02}	-	-	-	-	-	4,290	-

σ^2_u: additive genetic variance; σ^2_c: maternal environmental variance; σ_{uc}: additive genetic-maternal environmental covariance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; c^2: maternal environmental variance; r_{uc}: additive genetic-maternal environmental correlation; logL: natural log-likelihood; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; cAIC/DIC: conditional Akaike Information Criterion/Deviance Information Criterion; ρ/p_D: effective degrees of freedom/effective number of parameters; CI: confidence/credibility intervals in curly brackets

Table 2.2 Mean, standard deviation and range [minimum, maximum] of the true values and of the estimators of the simulated trait across models (M_1, M_2 and M_4) and methods of analyses (ASREML, WinBUGS and AnimalINLA) under two scenarios (low and high) for the direct genetic-maternal environmental correlation.

Model		M_1				M_2					M_4
Software	Simulated data	ASREML		AnimalINLA		ASREML		AnimalINLA		WinBUGS	
Scenario		low	high	low	high	low	high	low	high	high	
σ^2_u	7 (0.6) [6, 8]	15 (3) [11, 23]	19 (4) [13, 30]	18 (17) [13, 55]	32 (17) [13, 65]	6 (2) [4, 11]	10 (4) [4, 21]	14 (10) [11, 47]	26 (14) [14, 50]	19 (5) [7, 30]	10 (5) [5, 21]
σ^2_c	3 (0.5) [2, 4]	-	-	-	-	3 (1) [2, 6]	3 (1) [1, 7]	0	0	0.9 (0.8) [0, 3]	6 (3) [2, 13]
σ^2_e	32 (0.9) [30,34]	28 (2) [24, 30]	24 (2) [19, 29]	28 (2) [26, 31]	25 (2) [19, 29]	32 (2) [28, 35]	29 (2) [23, 34]	29 (2) [26, 31]	25 (2) [19, 29]	24 (3) [13, 30]	28 (6) [16, 35]
σ^2_p	42 (1.4) [39,46]	43 (2) [40, 48]	43 (2) [39, 48]	47 (17) [42, 84]	57 (17) [40, 94]	42 (2) [39, 46]	42 (2) [39, 46]	44 (10) [39, 78]	51 (13) [40, 78]	44 (3) [39, 51]	40 (6) [23, 45]
h^2	0.17 (0.02) [0.14,0.19]	0.35 (0.05) [0.27,0.47]	0.44 (0.07) [0.31,0.61]	0.44 (0.13) [0.30,0.65]	0.51 (0.14) [0.33,0.69]	0.15 (0.05) [0.08,0.26]	0.21 (0.09) [0.09,0.47]	0.34 (0.09) [0.26,0.60]	0.47 (0.13) [0.21,0.64]	0.43(0.11) [0.17,0.68]	0.24 (0.13) [0.13,0.54]
c^2	0.07 (0.01) [0.05,0.09]	-	-	-	-	0.07 (0.02) [0.04,0.13]	0.08 (0.03) [0.02,0.16]	0	0	0.02(0.02) [0, 0.09]	0.16 (0.09) [0.05,0.37]
σ_{uc}	-3.16 (0.47) [-4.29, -2.71]	-	-	-	-	-	-	-	-	-	-4.54 (4.62) [-9.74, -1.61]
σ_{uc}/σ^2_p	-0.08 (0.01) [-0.10, -0.06]	-	-	-	-	-	-	-	-	-	-0.13 (0.09) [-0.28, -0.04]
r_{uc}	-0.2/-0.8	-	-	-	-	-	-	-	-	-	-0.59 (0.2) [-0.94,-0.2]

σ^2_u: additive genetic variance; σ^2_c: maternal environmental variance; σ^2_e: residual variance; σ^2_p: phenotypic variance; h^2: direct heritability; c^2: maternal environmental variance; σ_{uc} = direct genetic-maternal environmental covariance; r_{uc}: direct genetic-maternal environmental correlation; N = 30 samples per scenario

Table 2.3 Mean Squared Errors of the variance components and the genetic parameters across models (M_1, M_2 and M_4) and methods of analyses (ASREML, WinBUGS and AnimalINLA) under two scenarios (low and high) of direct genetic-maternal environmental correlation.

Model	M_1				M_2					M_4
Software	ASREML		AnimalINLA		ASREML		AnimalINLA		WinBUGS	
Scenario	low	high	low	high	low	high	low	high	high	
σ_u^2	85.00	184.43	171.80	343.28	12.83	40.47	85.68	323.46	168.36	41.76
σ_c^2	-	-	-	-	2.60	4.67	9.00	9.00	6.05	20.88
σ_e^2	22.43	65.70	17.23	65.53	5.67	19.70	15.33	65.33	72.79	215.21
σ_p^2	6.99	12.36	177.28	199.08	5.78	7.30	45.24	182.53	11.73	128.33
h^2	0.04	0.08	0.09	0.18	0.01	0.02	0.04	0.14	0.44	0.81
c^2	-	-	-	-	0.01	0.01	0.01	0.01	0.03	0.85
r_{uc}	-	-	-	-	-	-	-	-	-	2.49

σ_u^2 : additive genetic variance; σ_c^2 : maternal environmental variance; σ_e^2 : residual variance; σ_p^2 : phenotypic variance; h^2 : direct heritability; c^2 : maternal environmental variance; r_{uc} : direct genetic-maternal environmental correlation; N = 30 samples per scenario

Table 2.4 Actual coverages of nominal 95% intervals of estimated variance components and genetic parameters as well as models percentage favoured by AIC (REML) and DIC (INLA) across models (M_1, M_2 and M_4) and methods of analyses (ASREML, WinBUGS and AnimalINLA) under two scenarios (low and high) of direct genetic-maternal environmental correlation.

	low				high					
	ASREML		AnimalINLA		ASREML		AnimalINLA		WinBUGS	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_2	M_4
σ_u^2	36.67	83.33	33.33	76.67	16.67	50.00	20.00	40.00	40.00	76.67
σ_c^2	-	86.67	-	-	-	56.67	-	-	63.33	93.33
σ_e^2	73.33	93.33	53.33	80.00	26.67	76.67	46.67	67.67	46.67	66.67
σ_p^2	80.00	96.67	73.33	90.00	70.00	86.67	66.67	80.00	86.67	86.67
h^2	33.33	76.67	33.33	73.33	13.33	53.33	20.00	33.33	36.67	63.33
c^2	-	90.00	-	-	-	56.67	-	-	60.00	93.33
r_{uc}	-	-	-	-	-	-	-	-	-	90.00
AIC/DIC	40.00	60.00	73.33	26.67	66.67	33.33	60.00	40.00	23.33	76.67

σ_u^2: additive genetic variance; σ_c^2: maternal environmental variance; σ_e^2: residual variance; σ_p^2: phenotypic variance; h^2: direct heritability; c^2: maternal environmental variance; r_{uc}: direct genetic-maternal environmental correlation; N = 30 samples per scenario

Table 2.5 Estimates of variance components and genetic parameters for the binary transformed body weight (BW) of broiler chicken at 35 days of age using a logit animal threshold model under three different methods software (ASREML, WinBUGS and AnimalINLA).

Software		σ^2_u	σ^2_p / σ^2_p*	h^2
ASREML (obs)	Mean (SE)	0.011 (0.003)	0.109 (0.003)	0.10 (0.02)
	CI (95%)	{0.006, 0.018}	{0.10, 0.12}	{0.04, 0.16}
ASREML	Mean (SE)	0.769 (0.226)	4.059 (0.226)	0.19 (0.05)
	CI (95%)	{0.34, 1.21}	{3.63, 4.49}	{0.09, 0.29}
WinBUGS	Mean (SE)	1.972 (0.859)	5.275 (0.795)	0.36 (0.09)
	CI (95%)	{0.87, 4.12}	{4.14, 7.27}	{0.21, 0.56}
AnimalINLA	Mean (SE)	0.866 (0.241)	4.156 (0.353)	0.21 (0.07)
	CI (95%)	{0.48, 1.41}	{3.77, 4.70}	{0.13, 0.30}

σ^2_u: additive genetic variance; σ^2_p: phenotypic variance; $\sigma^2_p* = \sigma^2_u + 3.29$ the phenotypic equivalent variance on the underlying scale; h^2: direct heritability; obs: observed scale; CI: confidence/credibility intervals in curly brackets

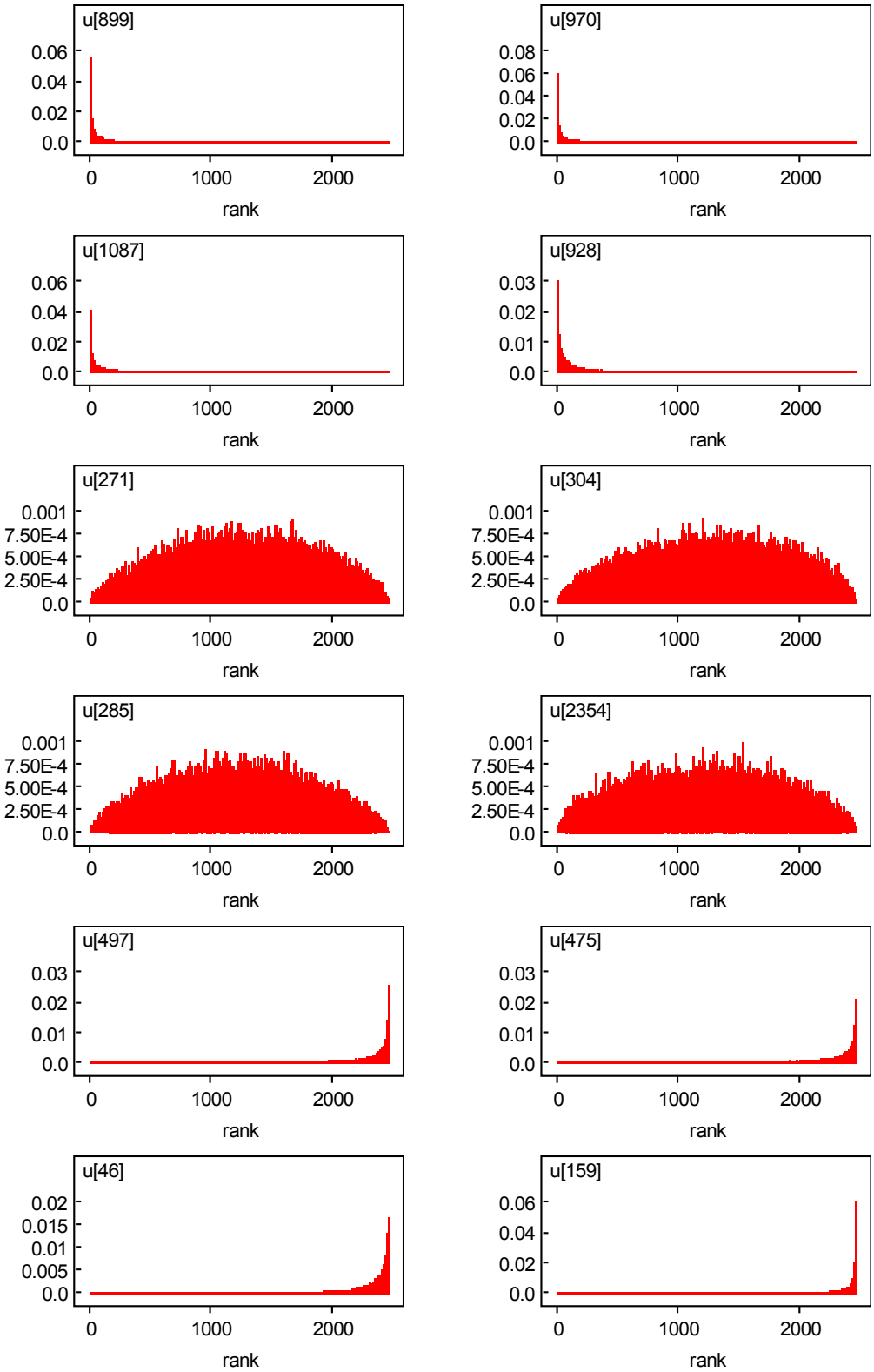

Figure 2.1 Distribution of ranking for twelve representative animals, based on the EBVs

3. Model comparison and estimation of genetic parameters for body weight in commercial broilers

3.1 Summary

The availability of powerful computing and advances in algorithmic efficiency allow for the consideration of increasingly complex models. Consequently, the development and application of appropriate statistical procedures for model evaluation is becoming increasingly important. This paper is concerned with the application of an alternative model determination criterion (conditional Akaike Information Criterion, cAIC) in a large dataset comprising 203,323 body weights of broilers, pertaining to 7 (BW_7) and 35 (BW_{35}) days of age. Seven univariate and seven bivariate models were applied. Direct genetic, maternal genetic and c^2 effects were estimated via REML. The model evaluation criteria included conditional Akaike Information Criterion (cAIC), Bayesian Information Criterion (BIC) and the standard Akaike Information Criterion (henceforth marginal; mAIC). According to cAIC the best-fitting model included direct genetic, maternal genetic and c^2 effects. Maternal heritabilities were low (0.10 and 0.03) compared to the direct heritabilities (0.17 and 0.21), while c^2 was 0.05 and 0.04 for BW_7 and BW_{35}, respectively. BIC and mAIC favoured a model that additionally included a direct-maternal genetic covariance, resulting in highly negative direct-maternal genetic correlations (-0.47 and -0.64 for BW_7 and BW_{35}, respectively) and higher direct heritabilities (0.25 and 0.28 for BW_7 and BW_{35}, respectively). Results suggest that cAIC can select different animal models than mAIC and BIC, having potential implications on selection decisions.

3.2 Introduction

Juvenile body weight has traditionally been considered as a trait of fundamental importance for broiler breeding programs. Heritability estimates for the additive genetic effects from age 0 to 36 weeks vary from 0.10 to 0.64 (see Table 3.1 for a review). Apart from these genetic effects, studies have also been carried out in an attempt to quantify the importance of maternal genetic effects in poultry. In those studies maternal heritability varied from 0.02 to 0.24, depending on model and population (Table 3.1). Studies that consider a non-zero covariance between direct additive and maternal genetic effects suggest a range between –0.11 and –0.92 for this correlation (Table 3.1). Additional studies are required, preferably using large data sets, in order to elucidate the magnitude of direct-maternal genetic correlation in poultry.

The estimation of appropriate genetic parameters depends on the formulation of a suitable statistical model. For a given class of candidate models the goal of statistical analysis is to estimate the model parameters and select the most appropriate model out of these candidates. A variety of information criteria have been proposed for this purpose. In the context of animal breeding, information criteria such as the Akaike's Information Criterion (AIC) (Akaike 1973) and the Bayesian Information Criterion (BIC) (Schwarz 1978) have been widely used. Both criteria are based on the log-likelihood and the number of estimated parameters p; the latter are used to penalize for the complexity of the models. When the estimation of random effects (such as the breeding values) is of interest, the number of model parameters is unclear. Recent studies in the statistical literature (Hodges and Sargent 2001; Vaida and Blanchard 2005; Liang and Wu 2008; Greven and Kneib 2010) suggest that in the case of linear mixed models conditional Akaike Information Criterion (cAIC), as opposed to the standard AIC, termed

marginal Akaike Information Criterion (mAIC), is unbiased when the (co)variance matrix of the random effects is known.

The objectives of the present study were a) the application and comparison of mAIC, BIC and cAIC for the evaluation of different animal models, b) the estimation of (co)variance components and genetic parameters for body weight at two different ages in a large data set of a commercial line of broiler chicken and c) the investigation of potential implications of model usage on selection decisions.

3.3 Materials and methods

3.3.1 Data selection and definition of fixed effects

Data on body weight at 7 (BW_7) and 35 (BW_{35}) days of age from a broiler line were made available by Aviagen Ltd. Fixed effects considered in the analysis included the gender and the week of hatch of animals. Progeny originated from groups of contemporary parents, referred to as mating groups, which were fitted as fixed effects in the analysis to account for the differences in the genetic level of parents. Finally, the age of sires and dams was considered as fixed effect by building four classes (9 - 12 months), ensuring that each one has sufficiently large number of records.

There is evidence that even a fully specified animal model may suffer from confounding between genetic and environmental effects (Kruuk and Hadfield 2007). A preliminary analysis using the SAS (2009) software revealed such a confounding between dams and sex of the offspring, as indicated by a loss of degrees of freedom. This was due to small families in which full-sibs all belonged to a particular sex class. All these records were eliminated from the analysis. In addition, three editing criteria were applied: 1) every dam had to have at least 4 offspring with records, following the suggestion from Lynch and Walsh (1998) to achieve equal precision in the estimates of paternal half-sibs and full-sibs variance 2) every sire should have mated with at least 3 dams and 3) animals were required to have records for both traits.

The final dataset consisted of 203,323 records comprising 99,330 males and 103,993 females, in 275 hatch weeks from 2000 to 2005. The pedigree included a total of 205,415 animals including 980 sires and 7,870 dams with progeny. A preliminary analysis of variance showed that the statistically significant fixed effects ($P<0.05$) for both traits included hatch, sex, mating group (93 classes) and the age of the parents. Hence, these fixed effects were included in all models. Interactions of sire by (all) fixed effects were not statistically significant and were not included in the analyses.

3.3.2 Statistical analysis

3.3.2.1 Univariate analysis

Differences in male and female mean performance as well as phenotypic variances were substantial. Initially, a simple additive animal model was applied for both traits and each sex separately, but no statistically significant differences in heritabilities were detected. Furthermore, the genetic correlation between male and female body weight approximated unity, indicating that the two traits are not sex-linked. Therefore, in the following analysis, body weights from males and females were treated as one trait, having sex as fixed effect in the model.

Seven animal models were considered for body weight at 7 and 35 days. Model M_1 was a purely direct additive model, while model M_2 allowed for the inclusion of maternal environmental effects. Model M_3 included the direct genetic and the full-sib

family effects, in order to account for possible dominance effects. A maternal genetic effect was incorporated in model M_4 in addition to the direct additive genetic effects, assuming zero direct-maternal genetic covariance (σ_{um}). Model M_5 was as model M_4, but with non-zero σ_{um}. Models M_6 and M_7 corresponded to models M_4 and M_5, respectively, but also included maternal environmental effects. In summary, the models in matrix notation were as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \quad (M_1)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_c c} + \mathbf{e} \quad (M_2)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_{fs} fs} + \mathbf{e} \quad (M_3)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_m m} + \mathbf{e}, \text{ with } \mathrm{cov(u,m)}=0 \quad (M_4)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_m m} + \mathbf{e}, \text{ with } \mathrm{cov}(u,m) = \sigma_{um} A \quad (M_5)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_m m} + \mathbf{Z_c c} + \mathbf{e}, \text{ with } \mathrm{cov(u,m)}=0 \quad (M_6)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Z_m m} + \mathbf{Z_c c} + \mathbf{e}, \text{ with } \mathrm{cov}(u,m) = \sigma_{um} A \quad (M_7),$$

where $\mathbf{y}$ = n×1 vector of observations (n = number of records = 203,323), $\mathbf{b}$ = p×1 vector of fixed effects (p = number of fixed effects classes = 374), $\mathbf{u}$ = q×1 vector of direct additive genetic effects (q = number of additive effects = 205,415), $\mathbf{m}$ = d×1 vector of maternal genetic effects (d = total number of females = 105,847), $\mathbf{c}$ = k×1 vector of maternal environmental effects (k = number of dams with offspring= 7,870), $\mathbf{fs}$ = s×1 vector of full-sib families (s = number of full-sib families = 8,609), $\mathbf{e}$ = n×1 vector of residuals; $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{Z_m}$, $\mathbf{Z_c}$ and $\mathbf{Z_{fs}}$ denote the incidence matrices relating the observations to the corresponding fixed and random effects; and $\mathbf{A}$ the additive relationship matrix. The vector of direct and maternal genetic effects was assumed to follow the multivariate normal distribution:

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{m} \end{bmatrix} \sim N\left(\mathbf{0}_{q+d}, \mathbf{G} \otimes \mathbf{A}\right),$$

where $\mathbf{0}_N$ denotes a N×1 vector of 0s, $\otimes$ denotes the Kronecker product,

$$\mathbf{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{um} \\ \sigma_{um} & \sigma_m^2 \end{bmatrix}$$

is the 2×2 (co)variance matrix between direct and maternal genetic effects, σ_u^2 denotes to the direct genetic variance, σ_m^2 the maternal genetic variance and σ_{um} the direct-maternal genetic covariance. The maternal environmental effects were assumed to follow a normal distribution given by: $\mathbf{c} \sim N\left(\mathbf{0}_k, \sigma_c^2 I_k\right)$, where $\mathbf{I_k}$ is an identity matrix of order k and σ_c^2 the maternal environmental variance. The full-sib family effect was also assumed to follow the normal distribution: $\mathbf{fs} \sim N\left(\mathbf{0}_s, \sigma_{fs}^2 \mathbf{I}_s\right)$, where σ_{fs}^2 denotes the full-sib family variance. Finally, residuals for the two traits were assumed normal as follows: $\mathbf{e} \sim N\left(\mathbf{0}_n, \sigma_e^2 \mathbf{I_n}\right)$, where σ_e^2 is the residual variance.

3.3.2.2 Bivariate analysis

In addition to the univariate analyses, seven bivariate models were applied. The most complex of the bivariate models used (model M_7, for both traits) can be represented as follows:

$$\begin{bmatrix} \mathbf{y}_7 \\ \mathbf{y}_{35} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_7 & 0 \\ 0 & \mathbf{X}_{35} \end{bmatrix} \begin{bmatrix} b_7 \\ b_{35} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{u7} & 0 \\ 0 & \mathbf{Z}_{u35} \end{bmatrix} \begin{bmatrix} \mathbf{u}_7 \\ \mathbf{u}_{35} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{m7} & 0 \\ 0 & \mathbf{Z}_{m35} \end{bmatrix} \begin{bmatrix} \mathbf{m}_7 \\ \mathbf{m}_{35} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{c7} & 0 \\ 0 & \mathbf{Z}_{c35} \end{bmatrix} \begin{bmatrix} \mathbf{c}_7 \\ \mathbf{c}_{35} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_7 \\ \mathbf{e}_{35} \end{bmatrix},$$

where subscript 7 (35) pertains to BW_7 (BW_{35}). The vector of direct and maternal genetic effects was assumed to follow the multivariate normal distribution:

$$\begin{bmatrix} \mathbf{u}_7 \\ \mathbf{u}_{35} \\ \mathbf{m}_7 \\ \mathbf{m}_{35} \end{bmatrix} \sim N(\mathbf{0}_{4n}, \mathbf{G} \otimes \mathbf{A}),$$

where $\mathbf{0}_N$ denotes a N×1 vector of 0s, $\otimes$ denotes the Kronecker product, **A** is the numerator relationship matrix, and

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{u7} & \sigma_{u7u35} & \sigma_{u7m7} & \sigma_{u7m35} \\ \sigma_{u7u35} & \sigma^2_{u35} & \sigma_{u35m7} & \sigma_{u35m35} \\ \sigma_{u7m7} & \sigma_{u35m7} & \sigma^2_{m7} & \sigma_{m7m35} \\ \sigma_{u7m35} & \sigma_{u35m35} & \sigma_{m7m35} & \sigma^2_{m35} \end{bmatrix}$$

is the 4×4 (co)variance matrix between direct and maternal genetic effects. Maternal environmental effects for the two traits were assumed to be correlated within animals and independent between, following a multivariate normal distribution as follows:

$$\begin{bmatrix} \mathbf{c}_7 \\ \mathbf{c}_{35} \end{bmatrix} \sim N(\mathbf{0}_{2k}, \mathbf{C} \otimes \mathbf{I}_k),$$

where $\mathbf{I}_k$ is an identity matrix of order k and

$$\mathbf{C} = \begin{bmatrix} \sigma^2_{c7} & \sigma_{c7c35} \\ \sigma_{c7c35} & \sigma^2_{c35} \end{bmatrix}$$

is the 2×2 (co)variance matrix between maternal environmental effects. Finally, residuals for the two traits were assumed to be correlated within animals and independent between as:

$$\begin{bmatrix} \mathbf{e}_7 \\ \mathbf{e}_{35} \end{bmatrix} \sim N(\mathbf{0}_{2n}, \mathbf{R} \otimes \mathbf{I}_n),$$

where $\mathbf{I}_n$ is an identity matrix of order n and

$$\mathbf{R} = \begin{bmatrix} \sigma^2_{e7} & \sigma_{e7e35} \\ \sigma_{e7e35} & \sigma^2_{e35} \end{bmatrix}$$

is the 2×2 residual (co)variance matrix.

All analyses (including univariate) were carried out by the ASREML software (Gilmour et al 2009). Estimates of direct (h^2) and maternal (h^2_m) heritability as well as maternal environmental effects (c^2) were calculated as the ratios of estimates of direct additive (σ^2_u), maternal genetic (σ^2_m) and maternal environmental (σ^2_c) variances respectively, to phenotypic variance (σ^2_p). The phenotypic variance accounts for the sum of all variance components, according to the model. The direct-maternal genetic correlation (r_{um}) was computed as the ratio of the estimate of direct-maternal genetic covariance (σ_{um}) to the product of the square roots of estimates of σ^2_u and σ^2_m. In addition, Willham (1972) was followed in calculating the total heritability (H^2_T) for BW_7 and BW_{35} by:

$$H^2_T = \frac{\sigma^2_u + 0.5\sigma^2_m + 1.5\sigma_{um}}{\sigma^2_p}.$$

3.3.2.3 Model evaluation criteria

Broadly speaking there are two classes of model evaluation and comparison procedures. When a true model exists and lies within the range of models entertained in the analysis, the BIC (Schwarz 1978) is consistent, in the sense that the true model will be selected by the BIC as more data accrue. The AIC gives an unbiased estimator of the Kullback-Leibler divergence of the current model from the true model. In addition, it can be shown that AIC dispenses with the need for a true model and chooses the model with the best short-term predictive ability (Stone 1977).

All model selection criteria require the computation of the deviance (D):

$$\mathrm{D} = -2\log(p(y|\hat{\theta})) = -2\log L,$$

where θ denotes the *px1* vector of the model parameters and $p(y|\hat{\theta})$ the likelihood of the data y evaluated at the maximum likelihood estimate $\hat{\theta}$. Akaike (1973) shows that the correct term for penalizing the deviance is twice the number of the model parameters *p*. Thus, he defined $AIC = -2\log L_i + 2p$ as the model selection criterion. A Bayesian argument was utilized by Schwarz (1978) to prove that the appropriate penalization term is *p*log (*n*) thus defining: $BIC = -2\log L_i + p\log n$, where n is the number of data observations.

The determination of the number of the model parameters is non-trivial when random effects are estimated using methods such as BLUP. For such cases the mAIC is shown in Crainiceanu and Ruppert (2004) to be asymptotically biased. In addition, Greven and Kneib (2010) showed that in linear mixed models mAIC is a biased estimator of the Akaike information due to the non-open parameter space and the lack of independence between observations. The conditional AIC (cAIC) defined by Vaida and Blanchard (2005) as $cAIC = -2\log L_i + 2\rho$ is asymptotically unbiased. A small correction term for estimating σ_u^2 is required, but Liang and Wu (2008) showed that cAIC represents an accurate approximation and ignoring the uncertainty arising from estimating the variance of the random effects has a small effect on its calculation. Notice that ρ, the effective degrees of freedom (Hodges and Sargent 2001), is given by the trace of the hat-matrix **H**, which is a linear map of observed to fitted values. In order to estimate ρ, the fourth field of ASREML's .yht file was utilized. This output corresponds to the diagonal elements of the so-called 'extended hat' matrix (Gilmour et al. 2009) and for the simple case where the residual variance is σ_e^2 I, the sum of these elements divided by the residual variance produce ρ. In the case of model M_1, it can be expressed as:

$$\rho = tr(\mathbf{H}) = tr\left(\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \lambda\mathbf{A}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{bmatrix}\right),$$

where $\lambda=\sigma_e^2/\sigma_u^2$. See the Appendix for a detailed description, including the case $\rho = 1 + (N-1)h^2$, when A=I. Minimizing the above criteria over a set of possible models can be seen as minimizing the average distance of an approximating model to the underlying truth (Greven and Kneib 2010). Thus, the model with the smallest cAIC value is to be preferred.

3.4 Results

3.4.1 Univariate analyses

Using the editing criteria, 401 outliers of BW_7 and 746 outliers of BW_{35} were traced and removed from the analysis. However, the estimated variance components were not affected indicating that the outliers had only a minimal contribution and impact on the dataset used. Table 3.2 summarizes the estimated variance components and genetic parameters of BW_7, along with likelihoods, ρ and model selection criteria for the seven univariate models. Direct heritability for BW_7 ranged from 0.17 to 0.49, while maternal heritability was between 0.10 and 0.25. Furthermore, maternal environmental variance accounted for 0.05-0.12 of the total phenotypic variance, depending on the model considered. A negative direct-maternal genetic correlation was detected ranging from -0.39 (model M_5) to -0.47 (model M_7). With regard to BW_{35} (Table 3.3), direct heritability varied from 0.21 to 0.43, while maternal heritability ranged from 0.03 to 0.11. Here, maternal environmental variance accounted for 0.04-0.05 of the total phenotypic variance, depending on the model. A negative direct-maternal genetic correlation was also estimated ranging from -0.35 (model M_5) to -0.64 (model M_7) for the trait. In both traits, the estimated variance of the full-sib families (model M_3) was equal to the maternal environmental variance (model M_2) indicating no significant importance of dominance effects.

The inclusion of maternal environmental effects (c^2) improved the fit significantly for both traits. The presence of maternal genetic (h_m^2) in addition to direct heritability (model M_4) improved the fit further compared to the purely additive animal model (M_1). Allowing for a non-zero direct-maternal genetic covariance (model M_7) also increased the log-likelihood, resulting in the smallest values for mAIC and BIC. However, this increase in log-likelihood was not significant when penalized using ρ via the cAIC. The model with the smallest ρ for both traits included direct genetic, maternal genetic and maternal environmental effects, with zero direct-maternal genetic covariance (model M_6). This reduction dominated the likelihood differences between the models and provided the best fit to the data based upon the cAIC.

According to cAIC, the best model included additive genetic, maternal genetic and maternal environmental effects while assuming no additive-maternal genetic covariance (model M_6). In this case, additive heritability was 0.17 and 0.21 for BW_7 (Table 3.2) and BW_{35} (Table 3.3), respectively. Maternal heritability was significantly higher in BW_7 (0.10) compared to BW_{35} (0.03), while maternal environmental variance accounted for 0.05 and 0.04 of the total phenotypic variance of the BW_7 and BW_{35}, respectively. In contrast, mAIC and BIC favored a model that incorporated direct genetic, maternal genetic and maternal environmental effects with a non-zero direct-maternal genetic covariance (model M_7). In this model, direct heritability was estimated to be 0.25 for BW_7 (Table 3.2) and 0.28 for BW_{35} (Table 3.3). Maternal heritability was found significantly higher in BW_7 (0.15) compared to BW_{35} (0.05), while maternal environmental variance accounted for 0.06 and 0.05 of the total phenotypic variance of BW_7 and BW_{35}, respectively. Furthermore, a negative additive-maternal genetic correlation was detected for BW_7 (-0.47), but was significantly higher for BW_{35} (-0.64).

Table 3.4 shows the rank (Spearman) correlations of animals based on the EBVs and the percentage of the best 1% males and 10% females that were common across the two best fitting models (M_6 and M_7). Although, high rank correlations were estimated among all animals for both traits (0.97 and 0.99 for BW_7 and BW_{35}, respectively), this

was not the case among the best males and females. With regards to BW_7, rank correlations were 0.72 and 0.79 among the best 1% males and 10% females, respectively. However, rank correlations of the best animals were considerably higher for BW_{35} for both sexes (0.89 and 0.94 for the best males and females, respectively). Slight effects of the model used were also reflected in the percentages of animals that were found to be the same across models. For BW_7, this percentage was 84% for the best males and 88% for the best females. For BW_{35}, the respective percentages were considerably higher reaching 92% and 94% for the best males and females, respectively.

3.4.2 Bivariate analysis

Bivariate analyses were conducted for all seven model combinations in order to estimate the genetic and the phenotypic correlation between the two traits. The three selection criteria proposed the same models used during the univariate analyses, while point estimates of all the parameters were not significantly different than those obtained from univariate analyses, for both traits. For all the above reasons, detailed results of the bivariate analyses are not presented, but the genetic, environmental and phenotypic correlations are given in Table 3.5. Under the application of a simple animal model the additive genetic and phenotypic correlations were estimated as high as 0.59 and 0.46, respectively. While the latter estimate ranged slightly between models (0.42-0.46), the additive genetic correlation varied considerably with values ranging from 0.11 to 0.59. The lowest estimates (0.11-0.18) for the genetic correlation were attained when maternal effects, either genetic or environmental, were included in the analysis. The correlations between the maternal genetic effects ranged slightly from 0.86 to 0.94 and this was also the case for the maternal environmental correlation (0.91 to 0.94).

3.5 Discussion

The present study demonstrates the application of a recently developed model selection criterion within the context of animal model(s). Results suggest that cAIC can select different models than the widely used mAIC and BIC. For random effects models the choice of the number of parameters incorporated in the penalty term (2p) is not always clear. When the number of variance components is included (mAIC), there have to be small differences in the log-likelihoods to result in the selection of simpler models. On the other extreme, if the number of parameters is considered to be equal to the total number of random effects, AIC tends to favor, on most occasions, the simplest model. Theoretical results suggest that cAIC is a more accurate model selection criterion when the random effects are included in the "focus" of the analysis (Vaida and Blanchard 2005; Liang and Wu 2008; Greven and Kneib 2010).

The cAIC is based upon the estimated effective degrees of freedom (Hodges and Sargent 2001) that can be approximated by the trace of the hat matrix. Because of the assumption that the estimated variances are in fact known, the true ρ might differ somewhat. An alternative to estimating ρ and cAIC would be the implementation of Bayesian methods for estimating the analogous effective number of parameters p_D. In order to compare ρ via REML, with the effective number of parameters (p_D) via MCMC methods a subset of approximately 60,000 records was used, spanning years 2002-2004. The results of this comparison (for both traits) for two univariate models (M_1 and M_2) showed that ρ and p_D did not differ significantly. A more detailed investigation of ρ is presented in the appendix.

A reduction of ρ was observed in model M_2 over model M_1 for BW_7 and this was more pronounced for BW_{35}. This could partly be attributed to the fact that ρ takes into account the information that is common to random effects through the additive relationship matrix A (see Appendix). However, this does not explain why smaller values were observed when maternal environmental effects were included for BW_7 or why there was such a big difference in models for BW_{35}. The ρ corresponds to the trace of the hat matrix, thus being dependent to the estimated variances and specifically $\lambda=\sigma_e^2/\sigma_u^2$. For BW_{35} the estimated σ_e^2 was much larger than σ_u^2, when compared to BW_7. Therefore, λ was much greater and this affected ρ for BW_{35}. Perhaps more importantly, the observed reduction in the effective degrees of freedom can be attributed to the negative correlation between direct and maternal genetic effects. Furthermore, this reduction of ρ in more complex models over the simple additive model might be an indication of the existence of a covariance between the random effects that is not adequately accounted for, such as a dam-offspring environmental covariance (Koerhuis and Thompson 1997).

All the criteria used for model comparison outlined the importance of the maternal genetic and maternal environmental effects while the inclusion of additive-maternal genetic covariance remained disputable. In contrast to the other two criteria (mAIC and BIC), cAIC did not suggest the inclusion of the latter covariance, which accounted for a reasonably significant part (0.08 – 0.09) of the total variance of the traits. Under the model favored by the cAIC criterion, the additive heritability estimate for BW_7 was smaller than the respective estimate for BW_{35} (0.17 vs. 0.21). Total heritability estimates for both traits were equal (0.22), indicating an equal efficiency of mass selection. Furthermore, estimates of maternal heritability were significantly higher for BW_7 than BW_{35} (0.10 vs. 0.03). This result may be reasonably expected as BW_7 is closer in terms of time to the egg weight. Yolk, albumen and egg weight have a large influence on chick weight, whereas the chick's own genes explain only a small part of its weight at hatching (Hartmann et al. 2003; Wolansky et al. 2004). Other studies (see Table 3.1) report additive and maternal heritability ranging from 0.10 to 0.61 and from 0.02 to 0.24, respectively. However, a wide range of weights, measured at later ages and models were considered in these studies making the reported estimates not directly comparable to ours. For body weight at 42 days of age, estimates of maternal heritability and c^2 effects did not differ from those reported by Koerhuis and Thompson (1997) but additive heritability was higher in the two populations studied on that occasion (0.23 and 0.27).

According to the other two criteria (BIC and mAIC) the additive heritability for both traits was of comparable magnitude (0.25 and 0.28 for BW_7 and BW_{35}, respectively), the maternal heritability was much higher in BW_7 (0.15) than in BW_{35} (0.05), while maternal environmental effects accounted for 0.06 and 0.05 of the total phenotypic variance of BW_7 and BW_{35}, respectively. The additive-maternal genetic correlation was estimated significantly higher in BW_{35} than in BW_7 (-0.64 vs. -0.47) indicating a more profound antagonism between the two effects at the 35 days of age. Here, estimates of BW_{35} did not differ from those reported by Koerhuis and Thompson (1997) for body weight at 42 days of age, with the exception of the additive-maternal genetic correlation which was found to be lower in magnitude (-0.54 vs. -0.64). It therefore seems that the use of the two standard criteria (mAIC and BIC) results in appreciably higher additive heritability estimates when contrasted to those suggested by the cAIC with implications on the expected selection response(s) for the two traits. Furthermore, the rather high additive-maternal genetic correlation estimated in the

present study for BW_{35} could not be attributed to sire by fixed effects interactions as suggested by other studies (e.g. Robinson 1996; Konstantinov and Brien 2003) since these interactions were of no importance here. Another possible explanation for inflated correlation might be the inability to account for other sources of (co)variance such as the dam-offspring environmental covariance. In a simulation study, Bijma (2006) has proposed that this covariance, when present and ignored, tends to inflate the estimated additive-maternal genetic correlation approximately equally to the value of the environmental correlation. Antagonism between direct and maternal genetic effects is expected to limit the potential for genetic improvement, palliating any response to selection on the offspring body weights. Furthermore, selecting for the body weight based on the animals' EBVs might result in opposite or even unfavorable directions for maternal attributes due to this negative correlation between additive and maternal genetic effects. However, such a negative covariance could be of importance in generating novel variation and in maintaining the genetic variance.

In an attempt to further quantify implications of model usage, such as on selection decisions, both the rank correlations of animals and the percentage of animals selected were calculated for the two best fitting models (M_6 and M_7). Based on these criteria, both models seemed to select the same animals with relatively small amount of losses for the breeding stock (6-16%). This was particularly true for BW_{35}, the main selection criterion. However, higher amount of loss of the best animals should be expected for BW_7 as suggested by the lower rank correlations for the trait.

Results of the bivariate analyses showed that under the application of a simple animal model (with additive genetic as the only random effects), the genetic correlation between the two traits was as high as 0.59. This result is in accordance with the genetic correlations in quails where values from 0.60 to 0.95 are reported (Saatci et al. 2003), depending on the correlated ages. However, significantly lower estimates (0.11-0.13) for this parameter were calculated when additional random effects of maternal origin, genetic or environmental, were included in the analysis. This result was rather unexpected considering that the two traits are recorded at only 28 days apart. To our knowledge, there are no genetic correlations of body weights at different ages reported in the literature that would allow for a direct comparison with our results. Relevant studies carried out in other species have shown that the additive genetic correlation between birth and weaning weight might range from 0.12 to 0.87 in sheep (Maria et al. 1993; Mousa et al. 1999; Hassen et al. 2003), from 0.64 to 0.89 in cattle (Tosh et al. 1999; Plasse et al. 2002), when maternal effects are accounted for, and in the region of 0.9 in turkeys for body weights measured in 14, 19 and 24 weeks of age (Kranis et al. 2006). A plausible explanation for this low genetic correlation relates this finding with the high correlations between maternal effects of both genetic and environmental origin. A significant amount of this positive covariance might thus be accounted for as additive. Such a scenario is however not fully supported in the literature, although most of the research regards different species (i.e. sheep, cattle) and not-directly comparable ages. When maternal (genetic and/or environmental) and direct additive genetic effects are simultaneously considered, there seems to be only a slight impact on the estimate of the additive genetic correlation (Maria et al. 1993; Mousa et al. 1999; Plasse et al. 2002; Hassen et al. 2003). Further studies in broilers are in need to comprehend the mechanism of maternal effects, the magnitude of additive-maternal genetic negative correlation and to elucidate the low additive genetic correlation between BW_7 and BW_{35} estimated here.

3.6 Appendix

3.6.1 A closer investigation of ρ

Three different pedigree examples of increasing relatedness were used for demonstration of the calculation of the effective degrees of freedom under the animal model, including additive genetic effects. The first pedigree (Pedigree A) consisted of 18 non-related animals. The additive relationship matrix A in this example is the identity matrix I. This is similar to the case examined in Hodges and Sargent (2001). In this example the effective degrees of freedom are expected to be close to:

$$\rho = tr(H) = tr\{\frac{1}{Nn}\frac{Nn(Nnh_e + h_u)}{nh_e + h_u} I_N\} = \frac{Nn + \frac{\sigma_e^2}{\sigma_u^2}}{n + \frac{\sigma_e^2}{\sigma_u^2}} = \frac{Nn + \lambda}{n + \lambda} = \frac{1 + (Nn - 1)h^2}{1 + (n - 1)h^2},$$

where $tr(H)$ is the trace of the hat-matrix, n the number of observations per animal, N the number of animals, $h_e = 1/\sigma_e^2$ and $h_u = 1/\sigma_u^2$. If n=1, i.e. one record per animal is assumed, then ρ would be: $\rho = 1 + (N - 1)h^2$.

In the second pedigree (Pedigree B), 9 non-inbred offspring resulted out of 9 base animals (Figure A1). The total number of animals is 18 like in the first example. Thus, the effect of incorporating the additive relationship matrix A on the computation of ρ was examined. Finally, a third pedigree (Pedigree C) was assumed, consisted of six base animals that gave birth to six offspring in the second generation. Offspring and parents were then mated (Figure A2) in order to examine the effect of close relationships and inbreeding on the computation of ρ. Following Vaida and Blanchard (2005), the penalizing term of the log-likelihood K_R in cAIC for a finite sample under REML method analysis is given as:

$$K_R = \frac{N - p - 1}{N - p - 2}(\hat{\rho} + 1) + \frac{p + 1}{N - p - 2},$$

where N the number of observations, p the number of classes of the fixed effects and $\hat{\rho}$ the estimated effective degrees of freedom. Note that K_R converges to $\hat{\rho}$, when N is much larger than p. An animal model was assumed. The equation of the model was: $y_i = \mu + u_i + e_i$, where y_i denotes the phenotypic value (observation), μ is the overall mean of the observations, u_i are the additive genetic values of animals and e_i denote the residual environmental effects associated with the i^{th} observation. For all three pedigree structures 50 different heritabilities were considered ranging from 0 to 1. Figure A3 depicts how ρ changes with h^2.

When h^2 tends to 1 (that is, when λ tends to 0) all three pedigrees exhibit the maximum degrees of freedom, which are equal to the total number of animals in the pedigree (N=18). This is to be expected since this extreme case implies infinite additive genetic variance and consequently genetically different animals. On the other hand, when σ_u^2 and h^2 converge to 0 (that is when λ approaches infinity), ρ tends to 1. This behavior is reasonable since $h^2 = 0$ suggests that no additive genetic heterogeneity exists. For all intermediate values of h^2, it seems that ρ decreases as the relations between individuals get closer and/or inbreeding increases (Figure A3). Specifically, for the same $h^2 \in (0,1)$, Pedigree C, that incorporates inbreeding and closer relations

between animals, has less effective degrees of freedom than Pedigree B, which in turn has smaller ρ compared to Pedigree A. This behavior appears intuitive and has been observed in a number of diverse but related contexts. For example, Kramer and Sugiyama (2011) report that ρ decreases as collinearity increases. In summary, it appears that ρ is appropriately adjusted through the additive relationship matrix A.

3.6.2 The comparison of ρ and p_D

The effective degrees of freedom (ρ) were contrasted to the comparable measure of the effective number of parameters (p_D) (Spiegelhalter et al. 2002). This was achieved via two univariate models (M_1 and M_2). For the fitting of the models via the R-package MCMCglmm (Hadfield 2010) a subset of 60,318 records was computationally feasible. The software used for the analyses were ASREML (Gilmour et al. 2009) and the R-package MCMCglmm (Hadfield 2010). The estimated p_D was derived under two different weakly-informative prior distributions: 1)σ_u, $\sigma_c \sim$ Uniform (0, 100) and 2) σ_u^2, $\sigma_c^2 \sim$ Inverse-Gamma (0.0001, 0.0001), whereσ_u,σ_c refer to the standard deviation of the direct genetic and the maternal environmental effects, respectively.

Table A1 shows that, although not identical, ρ and p_D do not differ significantly for both traits, regardless of the model applied. For a simple additive animal model (M_1) ρ was estimated to be 25563 and 16182, while p_D was estimated as 24971 and 16160 for BW_7 and BW_{35}, respectively. This was also the case for model M_2. In the model, which considered direct genetic and maternal environmental effects, ρ was estimated to be 23303 and 7145, while p_D was estimated at 22480 and 7135 for BW_7 and BW_{35}, respectively. In both cases, the relative difference ρ/p_D did not exceed 3%. The small differences between ρ and p_D could perhaps be attributed to the effect of the prior on the estimation of p_D and the fact that the variance components, estimated through REML, are implicitly assumed to be equal to the true values.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, Edition Petrov and Csaki, Budapest, Hungary.

Bijma, P. 2006. Estimating maternal genetic effects in livestock. *Journal of Animal Science* 84: 800-806.

Crainiceanu, C. and Ruppert, D. 2004. Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society, Series B* 66: 165–85.

Danbaro, G., Oyama, K., Mukai, F., Tsuji, S., Tateishi, T. and Mae, M. 1995. Heritabilities and genetic correlations from a selected experiment in broiler breeders using Restricted Maximum Likelihood. *Japanese Poultry Science* 32: 257-266.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. 2009. ASReml User Guide Release 3.0. *VSN International Ltd*, Hemel Hempstead, UK.

Greven, S. and Kneib, T. 2010. On the behaviour of marginal and conditional Akaike Information Criteria in Linear Mixed Models. *Biometrika* 97: 773-789.

Hadfield, J. 2010. MCMC methods for Multi-response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software* 33: 1-22.

Hartmann, C., Johansson, K., Strandberg, E. and Rydhmer, L. 2003. Genetic correlations between the maternal genetic effect on chick weight and the direct genetic effects on egg composition traits in a White Leghorn line. *Poultry Science* 82: 1-8.

Hassen, Y., Fuerst-Waltl, B. and Solkner, J. 2003. Genetic parameter estimates for birth weight, weaning weight and average daily gain in pure and crossbred sheep in Ethiopia. *Journal of Animal Breeding and Genetics* 120: 29–38.

Hodges, J.S. and Sargent, D.J. 2001. Counting degrees of freedom in hierarchical and other richly-parameterized models. *Biometrika* 88: 367-379.

Koerhuis, A.N.M and McKay, J.C. 1996. Restricted maximum likelihood estimation of genetic parameters for egg production traits in relation to juvenile body weight in broiler chickens. *Livestock Production Science* 46: 117-127.

Koerhuis, A.N.M and Thompson, R. 1997. Models to estimate maternal effects for juvenile body weight in broiler chickens. *Genetics Selection Evolution* 29: 225-249.

Konstantinov, K.V and Brien, F.D. 2003. Influence of sire by year interactions on the direct-maternal genetic correlation for weaning weight of Western Australian Merino sheep. *Australian Journal of Agricultural Research* 54: 723-729.

Kramer, N. and Sugiyama, M. 2011. The Degrees of Freedom of Partial Least Squares Regression. *Journal of the American Statistical Association*, to appear.

Kranis, A., Hocking, P.M., Hill, W.G. and Woolliams, J.A. 2006. Genetic parameters for a heavy female turkey line: Impact of simultaneous selection for body weight and total egg number. *British Poultry Science* 47: 685-693.

Kruuk, E.B. and Hadfield, J.D. 2007. How to separate genetic and environmental causes of similarity between relatives. *Journal of Evolutionary Biology* 20: 1890-1903.

Le Bihan-Duval, E., Beaumont, C. and Colleau, J.J. 1997. Estimation of the genetic correlations between twisted legs and growth or conformation traits in broiler chickens. *Journal of Animal Breeding and Genetics* 114: 239-259.

Liang, H. and Wu, H. 2008. A note on conditional AIC for linear mixed-effects models. *Biometrika* 95: 773-778.

Lynch, M. and Walsh, B. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Maria, G.A., Boldban, K.G. and van Vleck, L.D. 1993. Estimates of variances due to direct and maternal effects for growth traits of Romanov sheep. *Journal of Animal Science* 71: 845-849.

Mignon-Grasteau, S., Beaumont, C., Le Bihan-Duval, E., Poivey, J.P., Rochambeau, H. and Ricard, F.H. 1999. Genetic parameters of growth curve parameters in male and female chicken. *British Poultry Science* 40: 44-51.

Mousa, E., van Vleck, D.L. and Leymaster, K.A. 1999. Genetic parameters for growth traits for a composite terminal sire breed of sheep. *Journal of Animal Science 77*: 1659-1665.

Norris, D. and Ngambi, J.W. 2006. Genetic parameter estimates for body weight in local Venda chickens. *Tropical Animal Health and Production* 38: 605-609.

Plasse, D., Verde, O., Fossi, H., Romero, R., Hoogesteijn, R., Bastidas, P. and Bastardo, J. 2002. (Co)variance components, genetic parameters and annual trends for calf weights in a pedigree Brahman herd under selection for three decades. *Journal of Animal Breeding and Genetics* 119: 141–153.

Robinson, D.L. 1996. Models which might explain negative correlations between direct and maternal genetic effects. *Livestock Production Science* 45: 111-122.

Saatci, M., Ap Dewi, I. and Aksoy, A.R. 2003. Application of REML procedure to estimate the genetic parameters of weekly liveweights in one-to-one sire and dam pedigree recorded Japanese Quail. *Journal of Animal Breeding and Genetics* 120: 23–28.

SAS. 2009. SAS/STAT User's Guide, Version 9.2, Cary, NC, USA.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 75(4): 583-639 (with discussion).

Stone, M. 1977. An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal of the Royal Statistical Society, Series B* 30(1): 44-47.

Tosh, J.J., Kemp, R.A. and Ward, D.R. 1999. Estimates of direct and maternal genetic parameters for weight traits and backfat thickness in a multibreed population of beef cattle. *Canadian Journal of Animal Science* 79: 433-439.

Vaida, F. and Blanchard, S. 2005. Conditional Akaike information for mixed-effects models. *Biometrika* 92: 351-370.

Willham, R.L. 1972. The role of maternal effects in animal breeding: III.Biometrical aspects of maternal effects in animals. *Journal of Animal Science* 35: 1288-1293.

Wolansky, N.J., Renema, R.A., Robinson, F.E., Carney, V.L. and Fancher, B.I. 2004. Relationships among egg characteristics, chick measurements and early growth traits in ten broiler breeder strains. *Poultry Science* 86: 1784-1792.

Wolc, A., White, I.M.S., Avendano, S. and Hill, W.G. 2009. Genetic variability in residual variation of body weight and conformation scores in broiler chickens. *Poultry Science* 88: 1156-1161.

Tables and Figures

Table 3.1 Heritability estimates h^2 of additive (u), maternal genetic (m) effects and correlation coefficient (r_{um}) between additive and maternal genetic effects for juvenile body weight (BW) of broiler chicken.

Trait	h_u^2	h_m^2	r_{um}	References
BW_{7w}	0.10-0.33			Danbaro et al. 1995.
BW_{30w}	0.14-0.34			
BW_{6w}	0.28			Koerhuis and McKay 1996.
BW_{6w}	0.18-0.50	0.02-0.13	(-0.11)-(-0.92)	Koerhuis and Thompson 1997.
BW_{35w}	0.61			Le Bihan-Duval et al. 1997.
$BW_{8w\text{-}M}$	0.39	0.24		Mignon-Grasteau et al. 1999.
$BW_{36w\text{-}M}$	0.61	0.11		
$BW_{8w\text{-}F}$	0.45	0.21		
$BW_{36w\text{-}F}$	0.64	0.08		
$BW_{34d\text{-}M}$	0.21			Wolc et al. 2009.
$BW_{34d\text{-}F}$	0.25			
BW_{0w}	0.36	0.16	-0.19	Norris and Ngambi 2006.
BW_{4w}	0.25	0.13	-0.18	
BW_{10w}	0.41	0	0	
BW_{21w}	0.22	0	0	

where F corresponds to females and M males; w or d age of broilers in weeks or days, respectively

Table 3.2 Estimates of variance components, genetic parameters, log-likelihoods and model selection criteria under seven univariate models for the body weight (g) of broilers at 7 (BW_7) days of age.

Model	σ^2_u	σ^2_m	σ_{um}	σ^2_c	σ^2_{fs}	σ^2_e	σ^2_p	h^2	h^2_m	$\frac{\lvert\sigma_{um}\rvert}{\sigma^2_p}$	r_{um}	c^2	$\frac{\sigma^2_{fs}}{\sigma^2_p}$	H^2_T	logL	mAIC	BIC	cAIC	ρ
M_1	216	-	-	-	-	222	438	0.49	-	-	-	-	-	0.49	-695,689	1,391,380	1,391,390	1,527,932	68,276
	(4)					(2)	(2)	(0.01)						(0.01)					
M_2	85	-	-	47	-	275	408	0.21	-	-	-	0.12	-	0.21	-694,208	1,388,420	1,388,440	1,453,206	32,393
	(4)			(1)		(2)	(2)	(0.01)				(0.01)		(0.01)					
M_3	85	-	-	-	47	275	407	0.21	-	-	-	-	0.12	0.21	-694,186	1,388,376	1,388,396	1,453,339	32,482
	(4)				(1)	(3)	(3)	(0.01)					(0.01)	(0.01)					
M_4	76	82	-	-	-	279	437	0.17	0.19	-	-	-	-	0.27	-694,062	1,388,128	1,388,148	1,446,824	29,348
	(5)	(2)				(2)	(3)	(0.01)	(0.01)					(0.01)					
M_5	110	111	-43	-	-	262	440	0.25	0.25	0.10	-0.39	-	-	0.23	-694,011	1,388,028	1,388,059	1,467,907	39,940
	(6)	(4)	(4)			(3)	(6)	(0.02)	(0.01)	(0.01)	(0.03)			(0.01)					
M_6	72	42	-	22	-	281	416	0.17	0.10	-	-	0.05	-	0.22	-693,790	1,387,586	1,387,617	**1,444,399**	**28,199**
	(4)	(3)		(2)		(2)	(3)	(0.01)	(0.01)			(0.01)		(0.01)					
M_7	103	64	-38	23	-	265	417	0.25	0.15	0.09	-0.47	0.06	-	0.19	**-693,742**	**1,387,492**	**1,387,533**	1,467,008	39,557
	(7)	(5)	(4)	(2)		(4)	(3)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)		(0.01)					

σ^2_u: direct additive genetic variance; σ^2_m: maternal genetic variance; σ_{um} = direct-maternal genetic covariance; σ^2_c: maternal environmental variance; σ^2_{fs}: full-sib variance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; h^2_m: maternal heritability; r_{um}: direct-maternal genetic correlation; c^2: maternal environmental variance as proportion of σ^2_p; H^2_T: total heritability; logL: natural log-likelihood; mAIC: marginal Akaike Information Criterion; BIC: Bayesian Information Criterion; cAIC: conditional Akaike Information Criterion; ρ: effective degrees of freedom; standard errors in parenthesis

Table 3.3 Estimates of variance components, genetic parameters, log-likelihoods and model selection criteria under seven univariate models for the body weight (g) of broilers at 35 (BW_{35}) days of age.

Model	σ^2_u	σ^2_m	σ_{um}	σ^2_c	σ^2_{fs}	σ^2_e	σ^2_p	h^2	h^2_m	$\frac{\lvert\sigma_{um}\rvert}{\sigma^2_p}$	r_{um}	c^2	$\frac{\sigma^2_{fs}}{\sigma^2_p}$	H^2_T	logL	mAIC	BIC	cAIC	ρ
M_1	21,178	-	-	-	-	27,826	49,004	0.43	-	-	-	-	-	0.43	-1,178,923	2,357,848	2,357,858	2,474,044	58,098
	(490)					(270)	(280)	(0.01)						(0.01)					
M_2	9,979	-	-	2,304	-	32,939	45,223	0.22	-	-	-	0.05	-	0.22	-1,178,564	2,357,132	2,357,152	2,419,212	31,040
	(490)			(110)		(270)	(240)	(0.01)				(0.01)		(0.01)					
M_3	9,771	-	-	-	2,376	32,981	45,127	0.22	-	-	-	-	0.05	0.22	-1,178,521	2,357,046	2,357,066	2,418,494	30,724
	(470)				(100)	(260)	(240)	(0.01)					(0.01)	(0.01)					
M_4	11,320	3,171	-	-	-	32,330	46,821	0.24	0.07	-	-	-	-	0.28	-1,178,616	2,357,236	2,357,256	2,425,640	34,202
	(540)	(170)				(290)	(260)	(0.01)	(0.01)					(0.01)					
M_5	13,621	5,069	-2,948	-	-	31,177	46,919	0.29	0.11	0.06	-0.35	-	-	0.25	-1,178,588	2,357,182	2,357,213	2,437,672	40,245
	(790)	(320)	(420)			(410)	(270)	(0.02)	(0.01)	(0.01)	(0.04)			(0.01)					
M_6	9,304	940	-	1,733	-	33,265	45,241	0.21	0.03	-	-	0.04	-	0.22	-1,178,074	2,356,154	2,356,185	**2,415,644**	**29,294**
	(490)	(160)		(130)		(270)	(240)	(0.01)	(0.01)			(0.01)		(0.01)					
M_7	12,462	2,305	-3,440	2,189	-	31,683	45,198	0.28	0.05	0.08	-0.64	0.05	-	0.19	**-1,178,017**	**2,356,042**	**2,356,083**	2,432,611	37,836
	(760)	(270)	(380)	(130)		(390)	(260)	(0.02)	(0.01)	(0.01)	(0.04)	(0.01)		(0.01)					

σ^2_u: direct additive genetic variance; σ^2_m: maternal genetic variance; σ_{um} = direct-maternal genetic covariance; σ^2_c: maternal environmental variance; σ^2_{fs}: full-sib variance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; h^2_m: maternal heritability; r_{um}: direct-maternal genetic correlation; c^2: maternal environmental variance as proportion of σ^2_p; H^2_T: total heritability; logL: natural log-likelihood; mAIC: marginal Akaike Information Criterion; BIC: Bayesian Information Criterion; cAIC: conditional Akaike Information Criterion; ρ: effective degrees of freedom; standard errors in parenthesis

Table 3.4 Spearman correlation coefficients of EBVs (above the diagonal) and percentage of the best 1% males (m) and 10% females common (below the diagonal) under two univariate models (M_6 and M_7) of analysis, for body weight (g) of broilers at 7 (BW_7) and 35 (BW_{35}) days of age.

Model	Trait	**M_6**	M_7	Trait	**M_6**	M_7	Trait	**M_6**	M_7
M_6	BW_7	1	0.97	BW_{7m}	1	0.72	BW_{7f}	1	0.79
M_7			1		0.84	1		0.88	1
M_6	BW_{35}	1	0.99	BW_{35m}	1	0.89	BW_{35f}	1	0.94
M_7			1		0.92	1		0.94	1

Table 3.5 Phenotypic correlation coefficient (r_p), direct genetic correlation coefficient (r_u), maternal genetic correlation coefficient (r_m), correlation coefficients between maternal environmental effects (r_c), full-sib family effects (r_{fs}) and residual correlation (r_e) for the body weight (g) of broilers at 7 (BW_7) and 35 (BW_{35}) days of age, under seven bivariate models of analysis.

Models	r_u	r_m	r_c	r_{fs}	r_e	r_p
M_1	0.59 (0.01)	-	-	-	0.38 (0.01)	0.46 (0.01)
M_2	0.17 (0.03)	-	0.91 (0.01)	-	0.47 (0.01)	0.43 (0.01)
M_3	0.16 (0.04)	-	-	0.90 (0.01)	0.47 (0.01)	0.43 (0.01)
M_4	0.18 (0.04)	0.94 (0.01)	-	-	0.47 (0.01)	0.45 (0.01)
M_5	0.17 (0.04)	0.90 (0.01)	-	-	0.49 (0.01)	0.44 (0.01)
M_6	0.11 (0.04)	0.92 (0.03)	0.94 (0.02)	-	0.47 (0.01)	0.43 (0.01)
M_7	0.13 (0.04)	0.86 (0.02)	0.92 (0.02)	-	0.49 (0.01)	0.42 (0.01)

standard errors in parenthesis

Table A1. Estimates of the effective degrees of freedom (ρ) and the effective number of parameters (p_D), under two univariate homogeneous models of analyses (M_1 and M_2) for the body weight of broiler chicken at 7 (BW_7) and 35 (BW_{35}) days of age.

Model	Trait	ρ	p_D
M_1	BW_7	25563	24971
	BW_{35}	16182	16160
M_2	BW_7	23303	22480
	BW_{35}	7145	7135

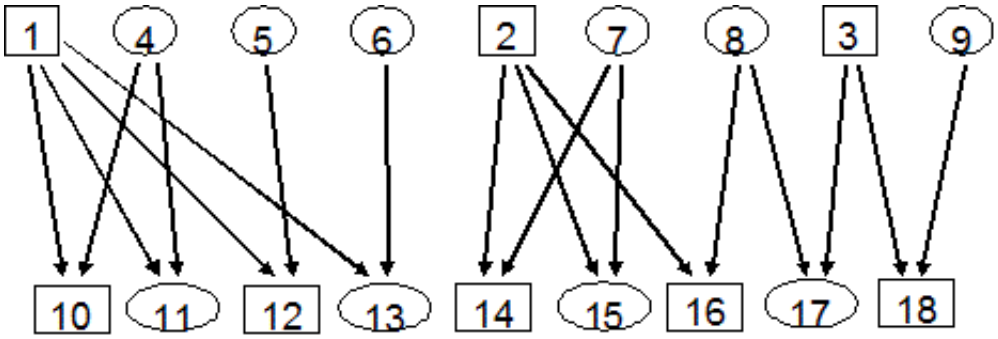

Figure A1. Pedigree B; where squares denote males and circles females

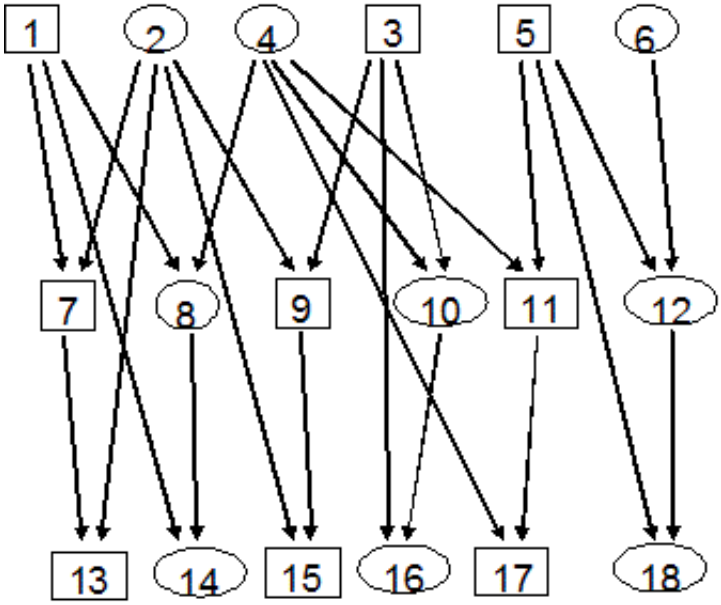

Figure A2. Pedigree C, where squares denote males and circles females

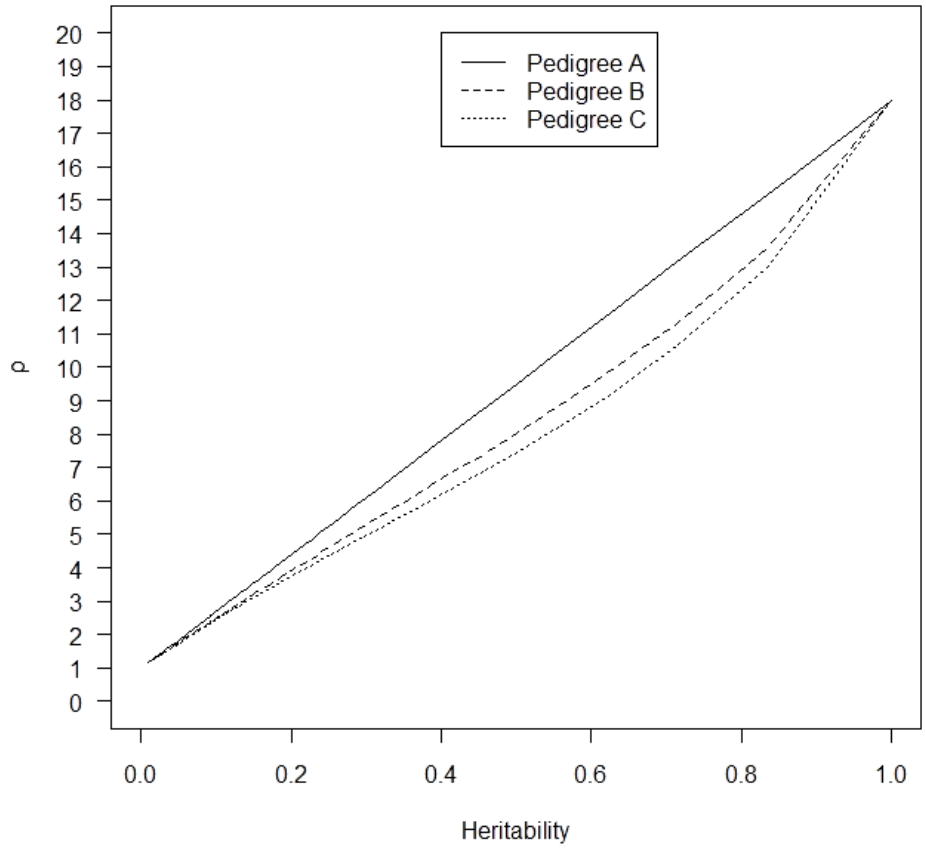

Figure A3. Effective degrees of freedom ρ under different levels of heritability h^2

4. Genetic analysis of sexual dimorphism of body weight in broilers

4.1 Summary

Variation in sexual dimorphism (SD) is particularly marked in meat type chickens. This paper investigates the genetic basis of SD in an important economic trait such as body weight at the 35 days of age (BW) in broilers by applying quantitative genetic analysis. A large dataset comprising 203,323 BW records of a commercial line of broiler chicken was used. First, a bivariate approach was employed treating BW as a sex-specific trait. During this approach, seven bivariate models were applied and variances due to direct additive genetic, maternal genetic and maternal environmental effects were estimated via Restricted Maximum Likelihood. The best-fitting model included direct additive genetic, maternal genetic and maternal environmental effects with a direct-maternal genetic covariance. Differences between male and female direct heritabilities were non-significant (0.28 vs. 0.29, males and females, respectively) implying no need for sex-specific selection strategies. The direct-maternal genetic correlation was more strongly negative in males than in females (-0.72 vs. -0.56), implying a more profound antagonism between direct additive and maternal genetic effects in this particular gender. The direct genetic correlation of BW between the two sexes was as high as 0.91 i.e. only slightly lower than unity. Second, variance components and genetic parameters of two measures of SD i.e. the weight difference (Δ) and the weight ratio (R) between the genders were estimated. Direct heritabilities for both measures were significantly different than 0 but of low magnitude (0.04). Apart from the additive-maternal covariance no other random effects were found to be of importance for Δ and R. Results of the present study suggest that only minimal selection responses due to selection of Δ and/or R and a small capacity for amplifying or reducing the BW differences between the sexes are to be expected, in this specific population. Furthermore, selection pressure on BW is expected to amplify sexual dimorphism.

4.2 Introduction

Sexual size dimorphism (SD), i.e. intersexual variation in phenotypic trait expression, is common throughout the animal kingdom. Sexual dimorphism is a key evolutionary feature that is related to ecology, behavior and life histories of organisms (Remes and Szekely 2010). Sexual dimorphism is commonly attributed to the combined effects of sex-specific selection pressure, sex-biased phenotypic and genetic variation and genetic correlations between sexes (Badyaev 2002) and is presumed to reflect adaptive divergence in response to selection favoring different optimal character states in the two sexes (Blackenhorn 2005). In comparisons among species, SD increases with body size in species where males are larger and decreases with body size in species where females are larger. This phenomenon is known as the Rensch's rule (Rensch 1950). One of the most distinctive examples of SD in animal kingdom is body weight (BW) in domestic chickens which is consistent with the Rensch's rule (males 21.5% heavier than females; Remes and Szekely 2010). The extent of SD does not differ among breed categories (cock fighting, ornamental and breeds selected for egg and meat production). Although SD of chicken breeds is not different from wild pheasants and allies, the wild ancestor of chickens (the red jungle fowl *G.gallus)* displays a more extreme SD (male 68.8% heavier) than any domesticated breed (Remes and Szekely 2010).

In meat type chickens i.e. broilers, juvenile BW is reported to be under the control of both direct and maternal genetic effects (Koerhuis and Thompson 1997; Mignon-Grasteau et al. 1998). There is also some evidence that the covariance between direct and maternal genetic effects might be of importance as well (Koerhuis and Thompson 1997). Sex-specific genetic analyses in broilers are scarce in the literature, while the direct-maternal genetic covariance is usually omitted. Mignon-Grasteau et al. (1998) have shown that both direct and maternal heritabilities of BW were significantly different between sexes in chicken. In that study, the direct-maternal genetic covariance was ignored. The covariance in question is omitted from analyses because of restrictions arising from limited data size and/or pedigree structures that allow for the application of only simplified genetic models.

One way of investigating the importance of the various genetic effects on SD is the application of bivariate models by treating measures on males and females as distinct traits. If the correlation between the two traits is significantly different than unity, then a sex-specific analysis may be justified. This approach takes into account any differences between the sexes and is useful when the primary target is the trait itself (Falconer and Mackay 1996). A second way of analyzing SD is the combination of measures taken on both sexes i.e. constructing an artificial trait. A common measure here is the difference in weight between males and females, hereafter called sexual difference and denoted Δ. Eisen and Legates (1966) were the first to develop an expression for the heritability of Δ, obtained from full-sib family means. An alternative expression is the ratio (R) of male to female body weight (Mignon-Grasteau et al. 1998). This parameter has the advantage of being scale free but it clearly shows some departure from normal distribution. Genetic analysis of these combined traits has the advantage of targeting the sex differences directly since selection based on these traits may amplify or reduce the sex differences. Under selection for inter-sexual uniformity, correlated responses such as sex-specific BWs are becoming of increasing interest for both from an evolutionary and a breeding point of view.

Given that the genetic basis of SD in broilers remains rather elusive, we have conducted the present study with the aim to provide answers to the following questions: a) is there sexual differentiation of direct and maternal genetic effects on BW in broilers, b) is the covariance between the direct and maternal effects of importance and possibly different between the two sexes, c) are two typical measures of SD (sex difference and sex ratio) under genetic control and d) what are the expected correlated responses for BWs when selecting for reduced SD?

4.3 Materials and methods

4.3.1 Data description, trait definition and fixed effect structure

Data on BW at 35 days of age from a male broiler line were made available by Aviagen Ltd. A preliminary analysis in SAS (2009) was performed to identify significant fixed effects. These included the week of hatch of animals (275 classes), spanning years 2003-2008, having on average 4 discrete generation equivalents. Progeny originated from groups of contemporary parents, referred to as mating groups (93 classes), which were also fitted as fixed effects in the analysis to account for the differences in the genetic level of parents. Finally, the age of sires and dams was considered as fixed effect by building four classes (9 - 12 months), ensuring that each one has sufficiently large number of records. When approximately equal precision is desirable in the estimates of paternal half-sibs variance and full-sibs variance, Lynch and Walsh (1998) suggested the allocation of at least 3 to 4 females per male and the

maintenance of full-sib families of approximately $1/(2\sigma_s^2/\sigma_p^2)$ progeny per female (but no less than 2), where σ_s^2 the sire variance and σ_p^2 the phenotypic variance. Thus, two editing criteria were applied: 1) every dam had to have at least 4 offspring with records and 2) every sire should mate with at least 3 dams. The final dataset consisted of 99,330 male and 103,993 female records. The pedigree included a total of 205,415 animals, of which 980 sires and 7,870 dams with progeny. The full-sib families consisted of 2 to 49 (average 12.81) males and 2 to 52 (average 13.48) females.

Assuming that both sexes appear in a litter, two artificial traits were constructed describing the sex difference (Δ) and the sex ratio (R) of male-female BW of randomly chosen pairs of full-sibs. Pairing of male and female records within full-sib families was made at random until either male or female records were exhausted. Surplus male or female records within a litter were omitted. Thus, 54,661 male-female pairs out of 6,995 full-sib families were made available for analysis. Table 4.1 shows descriptive statistics of all sexual dimorphism traits constructed. The sex difference (Δ) approximated a normal distribution and had an average of 275.4 g, while records regarding R were net numbers (average 1.133). Log transformation of R was used in an attempt to approximate normality and will refer as R hereafter. In order to account for a possible scale effect, an additional trait (called weighted sex difference and denoted wΔ) was also constructed by dividing Δ by the average BW of the full-sib pairs.

4.3.2 Bivariate analysis

Males were significantly heavier than females and differences in male and female phenotypic variances were also found to be substantial (Table 4.1). For this reason, BW was considered as a sex-limited trait and male (BW_m) and female (BW_f) weights were separately analyzed. Seven bivariate animal models were applied here. Model M_1 was a purely direct additive model, while model M_2 allowed for the inclusion of c^2 effects. Model M_3 included the direct genetic and the full-sib family effects, in order to account for possible dominance effects. A maternal genetic effect was incorporated in model M_4 in addition to the direct additive genetic effects, assuming zero direct-maternal genetic covariance (σ_{um}). Model M_5 was as model M_4, but with non-zero σ_{um}. Models M_6 and M_7 corresponded to models M_4 and M_5, respectively, but also included maternal environmental effects. In summary, the models in matrix notation were as follows:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} \quad (M_1)$$

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{c1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{c2}\end{bmatrix}\begin{bmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} \quad (M_2)$$

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{fs1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{fs2}\end{bmatrix}\begin{bmatrix}\mathbf{fs}_1\\ \mathbf{fs}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} \quad (M_3)$$

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{m2}\end{bmatrix}\begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

$$\text{with } \mathrm{cov}(u,m)=0 \quad (M_4)$$

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{m2}\end{bmatrix}\begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

$$\text{with } \operatorname{cov}(u,m) = \sigma_{um}\mathbf{A} \quad (M_5)$$

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{u1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{u2}\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{m2}\end{bmatrix}\begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{c1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{c2}\end{bmatrix}\begin{bmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

$$\text{with } \operatorname{cov}(u,m)=0 \quad (M_6)$$

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{u1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{u2}\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{m2}\end{bmatrix}\begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{c1} & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_{c2}\end{bmatrix}\begin{bmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

$$\text{with } \operatorname{cov}(u,m) = \sigma_{um}\mathbf{A} \quad (M_7),$$

where subscript 1 (2) pertains to male (female) BW; $\mathbf{y}_{1(2)}$ = $n_{1(2)}\times 1$ vector of observations ($n_{1(2)}$ = number of male {female} records = 99,330 {103,993}), $\mathbf{b}_{1(2)}$ = p×1 vector of fixed effects (p = number of fixed effects classes = 372), $\mathbf{u}_{1(2)}$ = q×1 vector of direct additive genetic effects (q = number of additive effects = 205,415), $\mathbf{m}_{1(2)}$ = d×1 vector of maternal genetic effects (d = total number of females = 105,847), $\mathbf{c}_{1(2)}$ = k×1 vector of maternal environmental effects (k = number of dams with offspring = 7,870), $\mathbf{fs}_{1(2)}$= s×1 vector of full-sib families (s= number of full-sib families= 8,609), $\mathbf{e}_{1(2)}$ = $n_{1(2)}\times 1$ vector of residuals; $\mathbf{X}_{1(2)}$, $\mathbf{Z}_{1(2)}$, $\mathbf{Z}_{\mathbf{m}1(2)}$, $\mathbf{Z}_{\mathbf{c}1(2)}$ and $\mathbf{Z}_{\mathbf{fs}1(2)}$ denote the incidence matrices relating the observations to the corresponding fixed and random effects. The vector of direct and maternal genetic effects was assumed to follow the multivariate normal distribution:

$$\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\\ \mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} \sim N\left(\mathbf{0}_{2q+2d}, \mathbf{G}\otimes\mathbf{A}\right),$$

where $\mathbf{0}_N$ denotes a N×1 vector of 0s, $\otimes$ denotes the Kronecker product, **A** is the additive relationship matrix

$$\mathbf{G} = \begin{bmatrix}\sigma^2_{u1} & \sigma_{u1u2} & \sigma_{u1m1} & \sigma_{u1m2}\\ \sigma_{u1u2} & \sigma^2_{u2} & \sigma_{u2m1} & \sigma_{u2m2}\\ \sigma_{u1m1} & \sigma_{u2m1} & \sigma^2_{m1} & \sigma_{m1m2}\\ \sigma_{u1m2} & \sigma_{u2m2} & \sigma_{m1m2} & \sigma^2_{m2}\end{bmatrix}$$

is the 4×4 (co)variance matrix between direct and maternal genetic effects, $\sigma^2_{u1(2)}$ denotes to the direct genetic variance, σ_{u1u2} the direct genetic covariance, $\sigma^2_{m1(2)}$ the maternal genetic variance, σ_{m1m2} the maternal genetic covariance and $\sigma_{u1(2)m1(2)}$ the direct-maternal genetic covariance. Maternal environmental effects for the two traits were assumed to follow the multivariate normal distribution:

$$\begin{bmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{bmatrix} \sim N\left(\mathbf{0}_{2k}, \mathbf{C}\otimes\mathbf{I}_k\right),$$

where $\mathbf{I}_k$ is an identity matrix of order k,

$$\mathbf{C} = \begin{bmatrix}\sigma^2_{c1} & \sigma_{c1c2}\\ \sigma_{c1c2} & \sigma^2_{c2}\end{bmatrix}$$

is the 2×2 (co)variance matrix between maternal environmental effects, $\sigma^2_{c1(2)}$ the maternal environmental variance and σ_{c1c2} the maternal environmental covariance. The full-sib family effect was also assumed to follow the multivariate normal distribution:

$$\begin{bmatrix} \mathbf{fs_1} \\ \mathbf{fs_2} \end{bmatrix} \sim N(\mathbf{0_{2s}}, \mathbf{D} \otimes \mathbf{I_s}),$$

where $\mathbf{I_s}$ is an identity matrix of order s,

$$\mathbf{D} = \begin{bmatrix} \sigma^2_{fs1} & \sigma_{fs1fs2} \\ \sigma_{fs1fs2} & \sigma^2_{fs2} \end{bmatrix}$$

is the 2×2 (co)variance matrix between full-sib family effects, $\sigma^2_{fs1(2)}$ denotes the full-sib family variance and σ_{fs1fs2} the full-sib families covariance. Finally, residuals, regarding different animals (males-females), were assumed independent between the two traits:

$$\begin{bmatrix} \mathbf{e_1} \\ \mathbf{e_2} \end{bmatrix} \sim N(\mathbf{0_{2n}}, \mathbf{R} \otimes \mathbf{I_n}),$$

where $\mathbf{I_n}$ is an identity matrix of order n and

$$\mathbf{R} = \begin{bmatrix} \sigma^2_{e1} & 0 \\ 0 & \sigma^2_{e2} \end{bmatrix}$$

is the 2×2 residual (co)variance matrix, where $\sigma^2_{e1(2)}$ is the residual variance.

4.3.3 Sexual dimorphism traits

Given that the two artificial traits concern full-sibs pairs, animal models were substituted by sire and dam models. Eight mixed linear models were applied for each artificial trait (Δ, R and wΔ). Model $M_{1'}$ and $M_{2'}$ were a sire and a dam model, respectively. Models $M_{3'}$ and $M_{4'}$ were sire models that allowed for the inclusion of maternal environmental effects and full-sib family effects, respectively. A dam effect was incorporated in model $M_{5'}$ in addition to the sire effect, assuming zero sire-dam genetic covariance σ_{sd}. Model $M_{6'}$ was as model $M_{5'}$, but with non-zero σ_{sd}. Models $M_{7'}$ and $M_{8'}$ corresponded to models $M_{5'}$ and $M_{6'}$, respectively, but also included maternal environmental effects. The same fixed effects as with the bivariate analysis were applied. In summary, the models in matrix notation were as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{e} \quad (M_{1'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_d d} + \mathbf{e} \quad (M_{2'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_c c} + \mathbf{e} \quad (M_{3'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_{fs} fs} + \mathbf{e} \quad (M_{4'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_d d} + \mathbf{e}, \text{ with } \mathrm{cov}(s,d)=0 \quad (M_{5'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_d d} + \mathbf{e}, \text{ with } \mathrm{cov}(s,d) = \sigma_{sd}\mathbf{A} \quad (M_{6'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_d d} + \mathbf{Z_c c} + \mathbf{e}, \text{ with } \text{cov(s,d)}=0 \quad (M_{7'})$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_d d} + \mathbf{Z_c c} + \mathbf{e}, \text{ with } \text{cov(s,d)} = \sigma_{sd}\mathbf{A} \quad (M_{8'}),$$

where $\mathbf{y}$ = $n_1 \times 1$ vector of observations (n_1 = number of pair records = 54,661), $\mathbf{b}$ = $p \times 1$ vector of fixed effects (p = number of fixed effects classes = 372), $\mathbf{s}$ = $q_1 \times 1$ vector of sire effects (q_1 = number of sires = 963), $\mathbf{d}$ = $d_1 \times 1$ vector of dam effects (d_1 = number of dams = 6,615), $\mathbf{c}$ = $k_1 \times 1$ vector of maternal environmental effects (k_1 = number of dams with offspring = 6,615), $\mathbf{fs}$ = $s_1 \times 1$ vector of full-sib families (s_1 = number of full-sib families = 6,995), $\mathbf{e}$ = $n_1 \times 1$ vector of residuals; $\mathbf{X}$, $\mathbf{Z_s}$, $\mathbf{Z_m}$, $\mathbf{Z_c}$ and $\mathbf{Z_{fs}}$ denote the incidence matrices relating the observations to the corresponding fixed and random effects. To retrieve estimates of variance components, model components were decomposed as follows: $\sigma_s^2 = \frac{1}{4}\sigma_u^2$, $\sigma_d^2 = \frac{1}{4}\sigma_u^2 + \sigma_m^2 + \sigma_{um}$ and $\sigma_{sd} = \frac{1}{4}\sigma_u^2 + \frac{1}{2}\sigma_{um}$. Thus, it is derived that: $\sigma_u^2 = 4\sigma_s^2$, $\sigma_{um} = 2\sigma_{sd} - \frac{1}{2}\sigma_u^2$ and $\sigma_m^2 = \sigma_d^2 - 2\sigma_{sd} + \sigma_s^2$, where σ_s^2 denotes the sire variance, σ_d^2 the dam variance and σ_{sd} the sire-dam covariance.

Estimates of direct (h^2) and maternal (h_m^2) heritability as well as maternal environmental effects (c^2) were calculated as ratios of estimates of direct additive (σ_u^2), maternal genetic (σ_m^2) and maternal environmental (σ_c^2) variances, respectively to phenotypic variance (σ_p^2). The phenotypic variance accounts for the sum of all variance components, according to the model. The direct-maternal genetic correlation (r_{um}) was computed as the ratio of the estimate of direct-maternal genetic covariance (σ_{um}) to the product of the square roots of estimates of σ_u^2 and σ_m^2. In addition, Willham (1972) was followed in calculating the total heritability (H_T^2) for BW by:

$$H_T^2 = \frac{\sigma_u^2 + 0.5\sigma_m^2 + 1.5\sigma_{um}}{\sigma_p^2}.$$

The estimation of variance components in all models was carried out with ASREML 3.0 (Gilmour et al. 2009).

Bivariate analyses between the two sex weights and the SD measures were also applied using models $M_{1'}$-$M_{8'}$ in order to evaluate the genetic and phenotypic correlations among the traits. However, only the sire model $M_{1'}$ converged. On this occasion the genetic correlation $r_{G_{PS}}$ between the primary trait P (BW_m or BW_f) and the secondary trait S (Δ or R) was given by: $r_{G_{PS}} = \frac{\sigma_{s_P s_S}}{\sigma_{s_P}\sigma_{s_S}}$, where $\sigma_{s_P s_S}$ the covariance between the sire effects and σ_{s_P}, σ_{s_S} the respective standard deviations. Note that a quadravariate sire model that included all traits failed to converge.

Besides the direct estimation of the genetic parameters of the sexual dimorphism traits, certain approximations are suggested in the literature based on the variance component estimation of the bivariate analysis of the primary traits. The heritability of the sex difference (Δ) is approximated according to Hanrahan and Eisen (1973) by:

$$h^2(\Delta)^* = \frac{(1 - r_{u1u2})\, h_{u1} h_{u2}}{(1 - r_{p1p2})}.$$

The heritability of the ratio of male to female body weight (R) was approximated using Sutherland's formula (Sutherland, 1965):

$$h^2(R)^* = \frac{h_1^2 CV_1^2 + h_2^2 CV_2^2 - 2r_{u1u2}\, CV_1 CV_2 h_1 h_2}{CV_1^2 + CV_2^2 - 2r_{p1p2}\, CV_1 CV_2},$$

where $h^2_{1(2)}$ denotes the heritability of the male (female) trait, $CV_{1(2)}$ the phenotypic coefficient of variation for the male (female) trait, r_{u1u2} and r_{p1p2} the additive genetic and phenotypic correlations between traits in males and females. All the above approximation formulae were used to compare our findings.

4.3.4 Model evaluation criteria

Model evaluation and comparison was carried out via two criteria: the Akaike Information Criterion (AIC; Akaike 1973) and the Bayesian Information Criterion (BIC; Schwarz 1978). The AIC gives an unbiased estimator of the Kullback-Leibler divergence of the current model from the true model. In addition, it can be shown that AIC dispenses with the need for a true model and chooses the model with the best short-term predictive ability (Stone 1977). When a true model exists and lies within the range of models entertained in the analysis, the BIC (Schwarz 1978) is consistent, in the sense that the true model will be selected as more data accrue. All model selection criteria require the computation of the deviance (D): $D = -2\log(p(y|\hat{\theta})) = -2\log L$, where θ denotes the *px1* vector of the model parameters and $p(y|\hat{\theta})$ the likelihood of the data y evaluated at the maximum likelihood estimate $\hat{\theta}$. Akaike (1973) showed that the correct term for penalizing the deviance is twice the number of the model parameters *p*. Thus, he defined $AIC = -2\log L_i + 2p$ as the model selection criterion. A Bayesian argument was utilized by Schwarz (1978) to prove that the appropriate penalization term is *p*log (*n*) thus defining: $BIC = -2\log L_i + p\log n$, where n is the number of data observations.

4.4 Results

4.4.1 Body weights

Table 4.2 summarizes the estimated variance components and genetic parameters of males and females, along with likelihoods and the information criteria for the six out of seven bivariate models. Model M_5, which accounted for the direct genetic effects, maternal genetic effects and a non-zero direct-maternal genetic covariance, did not achieve convergence and is thus not presented. Across models, direct heritability (h^2) for males ranged from 0.19 to 0.45, while maternal heritability (h^2_m) was between 0.02 and 0.08. Furthermore, maternal environmental variance (c^2) accounted for 0.05-0.06 of the total phenotypic variance. In females, h^2 varied from 0.23 to 0.44, while h^2_m ranged from 0.02 to 0.07, depending on the model. Here, c^2 accounted for 0.04-0.05 of the total phenotypic variance.

According to both information criteria, the best model included direct genetic effects, maternal genetic effects, maternal environmental effects and a non-zero direct-maternal genetic covariance (model M_7). In this model, h^2 was comparable across sexes

(0.28 and 0.29, for males and females, respectively) and this was also the case for h^2_m (0.07 and 0.05 for males and females, respectively). The same applied for c^2 (0.06 and 0.04 for males and females, respectively). Notably, the estimated direct-maternal genetic correlation r_{um} was significantly ($p<0.05$) higher in males than in females (-0.72 vs. -0.56).

The genetic and phenotypic correlations between the two sexes are shown in Table 4.3. Under a simple animal model, the additive genetic and phenotypic correlation was estimated as high as 0.95 and 0.43, respectively. While the latter estimate varied considerably across models (from 0.25 to 0.43), the additive genetic correlation ranged slightly (0.91 - 0.95). The correlations between the maternal genetic effects ranged slightly (0.86 - 0.94) and this was also the case for the maternal environmental correlation (0.91 to 0.94). According to model M_7, which was proposed by both criteria, the additive genetic correlation was estimated 0.91, the maternal genetic correlation 0.93 and the correlation of maternal environmental effects 0.94. Note that during all analyses, the residual correlation was set to zero, as the residuals regarded different animals (males or females), thus resulting in a remarkably low estimation of the phenotypic correlation (0.28) between the two traits.

4.4.2 Sexual dimorphism traits

Tables 4.4 and 4.5 summarize the variance components and genetic parameters estimations of Δ and (log-transformed) R, respectively. Given that point estimates of wΔ were not significantly different than those for Δ and R, only detailed results for the first two traits (Δ and R) are presented. For both SD measures, h^2 was significantly different than 0 ($P<0.05$) and ranged from 0.04 to 0.10, across models. Maternal genetic effects were of minor importance and practically equal to null. In both cases, c^2 accounted for a very small amount (0.017-0.022) of the total phenotypic variance. In both SD measures, both criteria suggested a model that incorporated direct genetic and c^2 effects, with a non-zero covariance between them (model $M_{8'}$). In this model, additive genetic effects were found to be significant different from zero ($h^2=0.04$) while a strong negative r_{um} was detected for both SD measures (-0.70 and -0.63 for Δ and R, respectively). Finally, the approximated heritabilities of Δ and R, as suggested by Hanrahan and Eisen (1973) and Sutherland (1965), respectively, were also calculated and no significant differences were detected, when compared to our estimates.

4.4.3 Genetic and phenotypic correlations between BW and SD

The genetic (r_G) and phenotypic (r_P) correlations between BW and the two SD measures are presented on Table 4.6. The genetic and phenotypic correlations between the two SD measures were estimated extremely high (0.97 ± 0.01 and 0.98 ± 0.01, respectively). The genetic correlation between Δ and BW_m was positive and of low to medium magnitude (0.34). The respective r_G between Δ and BWf was negative and of low magnitude (-0.20). An even lower r_G was estimated between R-BW_m (0.11) that could be treated as zero, considering the standard error of the estimate. Here, r_G between the R-BW_f was higher (-0.32) than that with Δ. As reasonably expected, both SD measures were positively and negatively correlated with BW_m and BW_f, respectively.

4.5 Discussion

4.5.1 Body weights

The present study focused on the various genetic effects – maternal included – that might play an important role on SD of BW in broilers. When the focus was on sex

specific BWs, all the model evaluation criteria outlined the importance of the direct genetic, the maternal genetic, the maternal environmental effects as well as the direct-maternal genetic covariance. Koerhuis and Thompson (1997) and Ilska et al. (2011) have already reported the significance of maternal effects on BW in broilers and the existence of a significantly negative r_{um} for this trait. In the present study, estimates of direct heritabilities were of medium magnitude and not different between sexes indicating that the two BWs are under the control of the same genes. Such estimates verify the existence of abundant additive variance in both sexes warranting appreciable selection responses and no need for undertaking sex-specific selection strategies. In contrast to present findings, sex differences of h^2 estimates are reported in the literature, in both chicken (0.28 vs. 0.43) and Muscovy ducks (0.40 vs. 0.51) (Mignon-Grasteau et al. 1998). Furthermore, there was a tendency for lower h^2_m in females than males (0.05 vs. 0.07), as already observed in chicken and Muscovy ducks (Mignon-Grasteau et al. 1998), in quails (Aggrey and Cheng 1994) and in turkeys (Chapuis et al. 1996).

The direct genetic correlation between the two traits was only slightly lower than unity suggesting that the two traits could be practically treated as one. Therefore the two traits are fully inter-dependent and independent selection of sexes is not possible (Falconer and Mackay 1996). Lower genetic correlations are reported by Mignon-Grasteau et al. (1998) in chicken (0.84) and Muscovy ducks (0.85). A salient finding of the present study was the higher r_{um} observed in males than in females (-0.72 vs. -0.56) suggesting a more profound antagonism between the two effects in this particular gender. As the equation of total heritability (Willham 1972) suggests, this negative σ_{um} is expected to have a limiting effect on selection response for the trait. A most plausible explanation for this strong negative covariance is antagonistic pleiotropy (Roff 2002) i.e. genes having antagonistic pleiotropic effects on maternal performance and offspring traits. Extensive meta-analysis of weight traits in domestic animals (Wilson and Reale 2006) strongly support this hypothesis while underlining the role of antagonistic pleiotropy in maintaining genetic variance (both direct additive and maternal genetic) in a trait under selection (Kirkpatrick and Lande 1989; Rasanen and Kruuk 2007). It should be noted, however, strong negative r_{um} may be spuriously estimated as a result of particular features of the data structure (Maniatis and Pollott 2003) or by failure to account for various interactions such as sire by herd or sire by year (Robinson 1996; Lee and Pollak 1997). These should be, however, ruled out in the present study since we had enough dam-offspring pairs and no significant sire interactions with any other term were of importance.

4.5.2 Sexual dimorphism

While sex-specific BWs were found to be under the control of a plethora of genetic effects, this was not the case for the combined weights whether difference or ratio. Only additive genetic effects were different than zero here and these were of low magnitude (h^2=0.04). This finding suggests that modification of SD through selection might be possible, when attempting inter-sexual uniformity, but very slow genetic progress should be anticipated. Other studies have reported higher h^2 in chicken (0.08) and Muscory ducks (0.13) (Mignon-Grasteau et al. 1998). Results of the present study suggest that selection for improved i.e. reduced SD results in unbalanced responses for the sex-specific BWs. Given the genetic correlations estimated herein, reduced SD is associated with decreasing and increasing BW for males and females, respectively. These changes are, however, not balanced between the sexes since an appreciable reduction of SD would request higher increases of BW in females than decreases in males. In any case, the two SD measures are highly inter-correlated fully supporting the

hypothesis of a co-evolution of the two traits. In addition, the correlations estimated herein convey that selection for higher BW is expected to amplify SD with a clear divergence between the sexes. Why this is also the case for BW in broilers remains an interesting scientific issue. Numerous studies in evolutionary biology across various species have been carried out in attempts to answer this question and it becomes clear that animal breeders have many lessons to be taught by studying the respective literature.

In conclusion, direct and maternal heritabilities between the sexes pose no significant differences. Thus no sex specific selection strategies are warranted while BW between sexes can be treated as one trait. The low direct heritabilities of the two SD measures suggest only a minimal response to selection and a small capacity for amplifying SD in this specific population. Unless a restriction index is practiced, selection pressure on BW is expected to amplify SD leading to higher divergence between sexes.

REFERENCES

Aggrey, S.A. and Cheng, K.M. 1994. Animal model analysis and (co)variances for growth traits in Japanese quail. *Poultry Science* 73: 1822-1828.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, Edition Petrov and Csaki, Budapest, Hungary.

Badyaev, A.V. 2002. Male and Female Growth in Sexually Dimorphic Species: Harmony, Conflict, or Both? *Comments on Theoretical Biology* 7: 10-33.

Blackenhorn, W.U. 2005. Behavioral causes and consequences of sexual size dimorphism. *Ethology* 111: 977-1016.

Chapuis, H., Ducrocq, V., Tixier-Boichard, M. and Delabrose, Y. 1996. Multivariate restricted maximum likelihood estimation of genetic parameters for production traits in three selected strains. *Genetics Selection Evolution* 28: 197-215.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R. 2009. ASReml User Guide Release 3.0. *VSN International Ltd*, Hemel Hempstead, UK.

Falconer, D.S. and Mackay, T.F. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group, Essex, UK.

Hanrahan, J.R. and Eisen, E.J. 1973. Sexual dimorphism and direct and maternal genetic effects on body weight in mice. *Theoretical Applied Genetics* 43: 39-45.

Ilska, J.J., Kranis, A., Burt, D.W. and Wooliams, J. 2011. The variance composition of broiler traits estimated on a large commercial population. *7th European Symposium on Poultry Genetics*, Edinburgh, Scotland.

Kirkpatrick, M. and Lande, R. 1989. The evolution of maternal characters. *Evolution* 43: 485-503.

Koerhuis, A.N.M. and Thompson, R. 1997. Models to estimate maternal effects for juvenile body weight in broiler chickens. *Genetics Selection Evolution* 29: 225-249.

Lee, C. and Pollak, E.J. 1997. Relationship between Sire × Year Interactions and Direct-Maternal Genetic Correlation for Weaning Weight of Simmental Cattle. *Journal of Animal Science* 75: 68-75.

Lynch, M. and Walsh, B. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Maniatis, N. and Pollott, G.E. 2003. The impact of data structure on genetic (co)variance components of early growth in sheep, estimated using an animal model with maternal effects. *Journal of Animal Science* 81: 101-108.

Mignon-Grasteau, S., Beaumont, C., Poivey, J.-P. and de Rochambeau, H. 1998. Estimation of the genetic parameters of sexual dimorphism of body weight in 'label' chickens and Muscovy ducks. *Genetics Selection Evolution* 30: 481-491.

Muller, H.J. 1958. Evolution by mutation. *Bulletin of the American Mathematical Society* 64: 137-160.

Rasanen, K. and Kruuk, L.E.B. 2007. Maternal effects and evolution at ecological time scales. *Functional Ecology* 21: 408-421.

Remes, V. and Szekely, T. 2010. Domestic chickens defy Rensch's rule: sexual size dimorphism in chicken breeds. *Journal of Evolutionary Biology* 23: 2754–2759.

Rensch, B. 1950. Die Abhaengigkeit der relativen Sexualdifferenz von der Koerpergroesse. *Bonner Zoologische Beitraege* 1: 58–69.

Robinson, D.L. 1996. Models which might explain negative correlations between direct and maternal genetic effects. *Livestock Production Science* 45: 111-122.

Roff, D.A. 2002. Life history evolution. Sinauer, Sunderland, MA.

SAS. 2009. SAS/STAT User's Guide, Version 9.2. Cary, NC, USA.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464

Stone, M. 1977. An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal of the Royal Statistical Society, Series B* 30(1): 44-47.

Sutherland, T.M. 1965. The correlation between feed efficiency and rate of gain, a ratio and its denominator. *Biometrics* 21: 739-739.

Weber, K.E. 1990. Increased selection response in larger populations. I. Selection for wing-tip height in *Drosophila melanogaster* at three population size. *Genetics* 125: 579-584.

Willham, R.L. 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. *Journal of Animal Science* 35: 1288-1293.

Wilson, A.J. and Reale, D. 2006. Ontogeny of additive and maternal genetic effects: lessons from domestic mammals. The *American Naturalist* 167: 23-38.

Tables

Table 4.1 Distribution of the sex difference (Δ), the sex ratio (R), the logarithm of the sex ratio (logR) and the weighted sex difference (wΔ) for body weight at 35 days of age.

	min	Q_1	median	mean	Q_3	max	sd	sk	kurt
Δ	-1030	110	280	275.44	440	1580	259.37	0.010	0.876
R	0.552	1.050	1.125	1.133	1.204	2.618	0.135	0.982	4.793
logR	-0.594	0.049	0.118	0.118	0.186	0.894	0.116	0.059	0.951
wΔ	-0.577	0.049	0.118	0.118	0.185	0.962	0.115	0.047	1.773

min: minimum value; Q_1: first quartile (25%); Q_3: third quartile (75%); max: maximum value; sd: standard deviation; sk: skewness; kurt: kurtosis

Table 4.2 Estimates of variance components, genetic parameters, log-likelihoods and model selection criteria under six bivariate models for the body weight (g) of male (m) and female (f) broiler chicken at 35 days of age.

Model	Sex	σ^2_u	σ^2_m	σ_{um}	σ^2_c	σ^2_{fs}	σ^2_e	σ^2_p	h^2	h^2_m	$\lvert\sigma_{um}\rvert / \sigma^2_p$	r_{um}	c^2	$\sigma^2_{fs} / \sigma^2_p$	H^2_T	logL	AIC	BIC
M_1	m	22,884 (600)	-	-	-	-	28,229 (350)	51,114 (350)	0.45 (0.01)	-	-	-	-	-	0.45 (0.01)	-1,177,305	2,354,612	2,354,622
	f	20,368 (500)	-	-	-	-	25,913 (300)	46,280 (300)	0.44 (0.01)	-	-	-	-	-	0.44 (0.01)			
M_2	m	9,524 (500)	-	-	2,833 (150)	-	34,362 (300)	46,720 (300)	0.20 (0.01)	-	-	-	0.06 (0.01)	-	0.20 (0.01)	-1,176,907	2,353,818	2,353,838
	f	10,775 (500)	-	-	2,086 (100)	-	30,151 (300)	43,012 (300)	0.25 (0.01)	-	-	-	0.05 (0.01)	-	0.25 (0.01)			
M_3	m	9,427 (500)	-	-	-	2,878 (150)	34,347 (300)	46,652 (300)	0.20 (0.01)	-	-	-	-	0.06 (0.01)	0.20 (0.01)	-1,176,862	2,353,728	2,353,748
	f	10,411 (500)	-	-	-	2,215 (100)	30,252 (300)	42,877 (300)	0.24 (0.01)	-	-	-	-	0.05 (0.01)	0.25 (0.01)			
M_4	m	11,398 (600)	3,655 (200)	-	-	-	33,549 (360)	48,601 (320)	0.23 (0.01)	0.08 (0.01)	-	-	-	-	0.27 (0.01)	-1,176,970	2,353,944	2,353,964
	f	11,780 (600)	3,000 (200)	-	-	-	29,705 (330)	44,486 (300)	0.26 (0.01)	0.07 (0.01)	-	-	-	-	0.29 (0.01)			
M_6	m	9,064 (500)	835 (150)	-	2,286 (150)	-	34,581 (300)	46,767 (300)	0.19 (0.01)	0.02 (0.01)	-	-	0.05 (0.01)	-	0.20 (0.01)	-1,176,865	2,353,736	2,353,767
	f	9,955 (500)	990 (150)	-	1,504 (150)	-	30,547 (300)	42,996 (300)	0.23 (0.01)	0.02 (0.01)	-	-	0.04 (0.01)	-	0.24 (0.01)			
M_7	m	12,895 (800)	3,031 (400)	-4,509 (500)	2,719 (150)	-	32,649 (460)	46,785 (300)	0.28 (0.02)	0.07 (0.01)	0.10 (0.01)	-0.72 (0.04)	0.06 (0.01)	-	0.17 (0.01)	**-1,176,792**	**2,353,592**	**2,353,633**
	f	12,649 (800)	2,167 (300)	-2,927 (400)	1,866 (150)	-	29,205 (400)	42,960 (300)	0.29 (0.02)	0.05 (0.01)	0.07 (0.01)	-0.56 (0.03)	0.04 (0.01)	-	0.21 (0.01)			

σ^2_u: direct additive genetic variance; σ^2_m: maternal genetic variance; σ_{um} = direct-maternal genetic covariance; σ^2_c: maternal environmental variance; σ^2_{fs}: full-sib variance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; h^2_m: maternal heritability; r_{um}: direct-maternal genetic correlation; c^2: maternal environmental variance as proportion of σ^2_p; H^2_T: total heritability; logL: natural log-likelihood; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; standard errors in parenthesis;

Table 4.3 Phenotypic correlation coefficient (r_p), direct genetic correlation coefficient (r_u), maternal genetic correlation coefficient (r_m), correlation coefficients between maternal environmental effects (r_c) and between full-sib family effects (r_{fs}) for the body weight (g) of male and female broiler chicken at 35 days of age, under six bivariate models of analysis.

Models	r_u	r_m	r_c	r_{fs}	r_p
M_1	0.95 (0.01)	-	-	-	0.43 (0.01)
M_2	0.91 (0.01)	-	0.87 (0.02)	-	0.26 (0.01)
M_3	0.91 (0.01)	-	-	0.92 (0.02)	0.25 (0.01)
M_4	0.92 (0.01)	0.97 (0.01)	-	-	0.30 (0.01)
M_6	0.91 (0.01)	0.93 (0.03)	0.94 (0.03)	-	0.25 (0.01)
M_7	0.91 (0.01)	0.93 (0.03)	0.94 (0.03)	-	0.28 (0.01)

standard errors in parenthesis

Table 4.4 Estimates of variance components, genetic parameters, log-likelihoods and model selection criteria under eight univariate models for the sex difference (Δ) in body weight (g) of male-female fullsib broiler chicken at 35 days of age.

Model	σ^2_u	σ^2_m	σ_{um}	σ^2_c	σ^2_{fs}	σ^2_e	σ^2_p	h^2	h^2_m	$\frac{\lvert\sigma_{um}\rvert}{\sigma^2_p}$	r_{um}	c^2	$\frac{\sigma^2_{fs}}{\sigma^2_p}$	H^2_T	logL	AIC	BIC
$M_{1'}$	3831	-	-	-	-	65815	66773	0.06	-	-	-	-	-	0.06	-330,760	661,522	661,531
	(460)					(400)	(400)	(0.01)						(0.01)			
$M_{2'}$	6892	-	-	-	-	65180	66903	0.10	-	-	-	-	-	0.10	-330,790	661,582	661,591
	(700)					(400)	(400)	(0.01)						(0.01)			
$M_{3'}$	2856	-	-	1481	-	64531	66726	0.04	-	-	-	0.022	-	0.04	-330,711	661,426	661,444
	(440)			(170)		(400)	(400)	(0.01)				(0.003)		(0.01)			
$M_{4'}$	2756	-	-	-	1616	64416	66722	0.04	-	-	-	-	0.024	0.04	-330,710	661,424	661,442
	(440)				(170)	(400)	(400)	(0.01)					(0.003)	(0.01)			
$M_{5'}$	3609	1018	-	-	-	65018	66822	0.05	0.015	-	-	-	-	0.06	-330,719	661,442	661,460
	(350)	(150)				(400)	(400)	(0.01)	(0.002)					(0.01)			
$M_{6'}$	3145	1085	-888	-	-	65052	67163	0.05	0.016	0.013	-0.48	-	-	0.04	-330,717	661,440	661,467
	(440)	(280)	(300)			(400)	(440)	(0.01)	(0.004)	(0.005)	(0.10)			(0.01)			
$M_{7'}$	2802	364	-	1113	-	64548	66726	0.04	0.006	-	-	0.017	-	0.05	-330,704	661,414	661,441
	(440)	(130)		(200)		(400)	(400)	(0.01)	(0.002)			(0.003)		(0.01)			
$M_{8'}$	2757	473	-807	1129	-	64552	67011	0.04	0.007	0.012	-0.70	0.017	-	0.03	-330,696	**661,404**	**661,440**
	(430)	(200)	(260)	(200)		(400)	(430)	(0.01)	(0.003)	(0.004)	(0.10)	(0.003)		(0.01)			

σ^2_u: direct additive genetic variance; σ^2_m: maternal genetic variance; σ_{um} = direct-maternal genetic covariance; σ^2_c: maternal environmental variance; σ^2_{fs}: full-sib variance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; h^2_m: maternal heritability; r_{um}: direct-maternal genetic correlation; c^2: maternal environmental variance as proportion of σ^2_p; H^2_T: total heritability; logL: natural log-likelihood; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; standard errors in parenthesis

Table 4.5 Estimates of variance components (multiplied by 10^5), genetic parameters, log-likelihoods and model selection criteria under eight univariate models for the logarithm of the sex ratio (R) of body weight of male-female fullsib broiler chicken at 35 days of age.

Model	σ_u^2	σ_m^2	σ_{um}	σ_c^2	σ_{fs}^2	σ_e^2	σ_p^2	h^2	h_m^2	$\frac{\lvert\sigma_{um}\rvert}{\sigma_p^2}$	r_{um}	c^2	$\frac{\sigma_{fs}^2}{\sigma_p^2}$	H_T^2	logL	AIC	BIC
$M_{1'}$	68	-	-	-	-	1329	1346	0.05	-	-	-	-	-	0.05	-240,303	480,608	480,617
	(8)					(8)	(8)	(0.01)						(0.01)			
$M_{2'}$	122	-	-	-	-	1318	1348	0.09	-	-	-	-	-	0.09	-240,360	480,722	480,731
	(14)					(8)	(8)	(0.01)						(0.01)			
$M_{3'}$	52	-	-	27	-	1305	1345	0.04	-	-	-	0.020	-	0.04	-240,211	480,426	480,444
	(8)			(4)		(8)	(8)	(0.01)				(0.003)		(0.01)			
$M_{4'}$	50	-	-	-	29	1303	1345	0.04	-	-	-	-	0.022	0.04	-240,201	480,406	480,424
	(10)				(4)	(10)	(8)	(0.01)					(0.003)	(0.01)			
$M_{5'}$	65	18	-	-	-	1314	1347	0.05	0.014	-	-	-	-	0.05	-240,223	480,450	480,468
	(7)	(3)				(8)	(8)	(0.01)	(0.002)					(0.01)			
$M_{6'}$	57	15	-12	-	-	1315	1355	0.04	0.011	0.009	-0.40	-	-	0.03	-240,217	480,440	480,467
	(8)	(5)	(5)			(9)	(9)	(0.01)	(0.004)	(0.004)	(0.10)			(0.01)			
$M_{7'}$	51	9	-	19	-	1306	1345	0.04	0.006	-	-	0.014	-	0.04	-240,197	480,400	480,427
	(8)	(3)		(4)		(8)	(8)	(0.01)	(0.002)			(0.003)		(0.01)			
$M_{8'}$	50	7	-12	19	-	1306	1352	0.04	0.005	0.009	-0.63	0.014	-	0.03	-240,188	**480,384**	**480,420**
	(8)	(4)	(5)	(4)		(9)	(9)	(0.01)	(0.003)	(0.003)	(0.10)	(0.003)		(0.01)			

σ_u^2: direct additive genetic variance; σ_m^2: maternal genetic variance; σ_{um} = direct-maternal genetic covariance; σ_c^2: maternal environmental variance; σ_{fs}^2: full-sib variance; σ_e^2: residual variance; σ_p^2: phenotypic variance in g^2; h^2: direct heritability; h_m^2: maternal heritability; r_{um}: direct-maternal genetic correlation; c^2: maternal environmental variance as proportion of σ_p^2; H_T^2: total heritability; logL: natural log-likelihood; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; standard errors in parenthesis

Table 4.6 Genetic (r_G) and phenotypic (r_P) correlations between BW_m or BW_f and Δ or R.

	r_G		r_P	
	BW_m	BW_f	BW_m	BW_f
Δ	0.34 (0.06)	-0.20 (0.07)	0.54 (0.01)	-0.46 (0.01)
R	0.11 (0.06)	-0.32 (0.07)	0.45 (0.01)	-0.55 (0.01)

standard errors in parenthesis

5. Genotype (by sex) by environment interaction for body weight in broiler chicken

5.1 Summary

A study was conducted to evaluate the effect of genotype by environment interaction (GEI) on body weight (BW) at 35 days of age for commercial broilers. A large dataset was used, consisting of 203,323 and 35,595 records obtained in high (H) and low (L) hygiene conditions, respectively. Bi- and quadra-variate analyses of environmental and sex-environmental specific traits were employed in an attempt to estimate genetic parameters. BW in the two different environments was treated as two distinct traits (BW_H, BW_L) in the bi-variate analyses, while it was analyzed separately for each sex in each environment ($BW_{H♂}$, $BW_{H♀}$, $BW_{L♂}$ and $BW_{L♀}$) via quadra-variate analyses. Variance components due to direct additive genetic, maternal genetic and maternal environmental effects were estimated via Restricted Maximum Likelihood. Model fit was assessed by the conditional Akaike Information Criterion (cAIC), the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). The best fitting models that successfully converged always included the direct genetic and the maternal environmental effects. During bivariate analysis, the direct heritability of BW was significantly higher in the L environment when contrasted to H (0.27 vs. 0.14), while the maternal environmental variance accounted for 0.05 of the total phenotypic variance in both environments. During the quadravariate approach, the direct heritabilities were estimated as high as 0.14, 0.19, 0.26 and 0.30 for $BW_{H♂}$, $BW_{H♀}$ $BW_{L♂}$ and $BW_{L♀}$, respectively, whence the maternal environmental variance accounted for 0.06-0.08 of the total phenotypic variance of the trait(s). The direct genetic correlations between the two environments ranged from 0.28 to 0.45, indicating the presence of strong GEI. The direct genetic correlations between the two sexes within the same environment ($BW_{H♂}$-$BW_{H♀}$ and $BW_{L♂}$-$BW_{L♀}$) were of high magnitude (0.81). Based on the current findings, no sex-specific selection policies are to be pursued while GEI should be appropriately accounted for during genetic evaluation.

5.2 Introduction

Differential performance of specific genotypes under changing production conditions implies the presence of genotype by environment interaction (GEI). This interaction can reduce the accuracy of a model in predicting animal performance and may have undesirable effects on the estimation of breeding values (Case et al. 2010) leading to possible re-ranking of individuals between environments. As a result, selection in one environment will not necessarily lead to enhanced performance under different environmental conditions (Falconer and Mackay, 1996). Falconer and Mackay (1996) have suggested a method to quantify the importance of GEI by treating performances in two or more environments as different traits and estimating the genetic correlation (r_G) between them via a multivariate approach. However, it is not clear which value of r_G implies significant GEI. According to Robertson (1959), GEI is important if r_G between the environments is less than 0.80. Mulder et al. (2006) suggested that when r_G is lower than 0.70-0.50, then environment-specific breeding programs are necessary for special adaptability in environments of importance.

With the worldwide distribution of stocks, a major poultry breeding objective is breeding commercial broilers that perform adequately in a wide variety of environmental conditions. The capacity of single genotypes to exhibit a range of phenotypes in response to variation in the environment is defined as phenotypic

plasticity (Fordyce, 2006). GEI might involve different nutritional (Havenstein et al., 1994), management (N'Dri et al., 2007), climatic (Cahaner and Leenstra, 1992; Leenstra and Cahaner, 1992; Yalcin et al., 1997, Settar et al., 1999) or hygiene conditions (Banos et al., 2006; Ye et al., 2006; Long et al., 2008). The latter conditions might have a dramatic impact on the performance of broilers, but there seems to be minimal relevant information reported in the literature, at least from a quantitative genetics point of view. The studies of Ye et al. (2006) and Long et al. (2008) were mainly concerned with identifying immune-related genes or SNP subsets that are associated with the differential performance of broilers between different hygiene conditions. Given this lack of information the present study was conducted aiming at the following questions: a) do the hygiene conditions pose significant GEI for an economically important trait such as the body weight (BW) at 35 days of age in broilers and b) if so, are sex-specific breeding policies necessary to accommodate this interaction? In order to answer these questions we have used a large data set of BW obtained under two different hygiene level environments and followed both bi- and quadra-variate analyses.

5.3 Materials and method

5.3.1 Data description

Data on BW at 35 days of age were made available by Aviagen Ltd., regarding a commercial line of broiler chicken, raised in either a high (H) or low (L) hygiene condition environment. The high and low hygiene conditions were representative of selection nucleus and low hygiene commercial level conditions respectively. Information in both environments included sire, dam, age of sire and dam, hatch week, mating group and sex of each bird. Regarding the H environment, all broilers belonging to three weekly hatches were reared together. Potential parents were selected each week and allocated to mating groups. A mating group typically comprised approximately 8 sires, each mated to roughly 10 dams in a nested structure for a standard number of hatch weeks. Progeny born in a particular hatch from a particular mating group came from contemporaneous parents and shared the same environment. Every second hatch week, 290 offspring on average were randomly selected out of the full-sib families and were raised in L environment.

Table 5.1 presents aspects of the dataset and the pedigree used. The dataset consisted of 203,323 (99,330 male and 103,993 female) and 35,595 (17,990 male and 17,605 female) records in H and L environment, respectively. The pedigree included a total of 241,010 animals, of which 980 sires and 7,870 dams with progeny in H environment and 665 sires and 3,822 dams with offspring in L environment. All sires and dams with offspring in L environment also had offspring in H environment, thus providing the genetic links that make between-environment analysis possible. Sires produced 4 to 605 offspring (average 207.47) of which 2 to 300 (average 101.36) male and 2 to 337 (average 106.88) female offspring, all raised in H. Additionally, 4 to 179 (average 53.53) offspring per sire were raised in L environment. Dams produced 4 to 88 (average 25.84) offspring of which 2 to 49 (average 12.81) male and 2 to 52 (average 13.48) female offspring, all raised in H. In addition, 4 to 30 (average 9.33) offspring per dam were raised in L environment.

5.3.2 Statistical analysis

Broilers in H were significantly (P<0.01) heavier compared to the L environment (average 2,329 g vs. 1,514 g). In addition, males were significantly heavier than females in both environments (average 2,470.2 g vs. 2,194.7 g in H and 1,591.4 g vs 1,433.9 g in L). Such a description may imply the need of treating BW either as an environmental-specific or even as a sex by environment specific trait leading to bivariate and quadravariate analyses, respectively. In both cases a number of models were fitted including direct genetic, maternal genetic and/or maternal environmental effects. However, models that simultaneously considered maternal genetic and maternal environmental effects suffered from convergence issues. Indeed, preliminary univariate analysis of BW in the L environment revealed that maternal genetic and maternal environmental effects could not be sufficiently disentangled in the specific environment. This was probably due to the fact that the females in the L environment did not produce any offspring. Consequently, only four models were applied in both the bi- and quadra-variate analyses. With regard to the fixed effect part of the models, preliminary analysis of variance showed that the statistically significant (P<0.05) fixed effects for all traits included hatch (275 and 122 weeks in H and L, respectively), sex (where appropriate), mating group (93 and 42 classes in H and L, respectively) and the age of the parents (four classes from 9 to 12 months). The above fixed effects were included in all models.

5.3.2.1 Bivariate analysis

Four animal models were considered. Model B_1 was a purely direct additive model, while model B_2 allowed for the inclusion of maternal environmental effects. A maternal genetic effect was incorporated in model B_3 in addition to the direct additive genetic effects, assuming zero direct-maternal genetic covariance (σ_{um}). Model B_4 was as model B_3, but with non-zero σ_{um}. In matrix notation these models can be described as follows:

$$\begin{bmatrix}\mathbf{y}_H\\ \mathbf{y}_L\end{bmatrix} = \begin{bmatrix}\mathbf{X}_H & 0\\ 0 & \mathbf{X}_L\end{bmatrix}\begin{bmatrix}\mathbf{b}_H\\ \mathbf{b}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_H & 0\\ 0 & \mathbf{Z}_L\end{bmatrix}\begin{bmatrix}\mathbf{u}_H\\ \mathbf{u}_L\end{bmatrix} + \begin{bmatrix}\mathbf{e}_H\\ \mathbf{e}_L\end{bmatrix} \quad (B_1)$$

$$\begin{bmatrix}\mathbf{y}_H\\ \mathbf{y}_L\end{bmatrix} = \begin{bmatrix}\mathbf{X}_H & 0\\ 0 & \mathbf{X}_L\end{bmatrix}\begin{bmatrix}\mathbf{b}_H\\ \mathbf{b}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_H & 0\\ 0 & \mathbf{Z}_L\end{bmatrix}\begin{bmatrix}\mathbf{u}_H\\ \mathbf{u}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{\mathbf{c}_H} & 0\\ 0 & \mathbf{Z}_{\mathbf{c}_L}\end{bmatrix}\begin{bmatrix}\mathbf{c}_H\\ \mathbf{c}_L\end{bmatrix} + \begin{bmatrix}\mathbf{e}_H\\ e_L\end{bmatrix} \quad (B_2)$$

$$\begin{bmatrix}\mathbf{y}_H\\ \mathbf{y}_L\end{bmatrix} = \begin{bmatrix}\mathbf{X}_H & 0\\ 0 & \mathbf{X}_L\end{bmatrix}\begin{bmatrix}\mathbf{b}_H\\ \mathbf{b}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_H & 0\\ 0 & \mathbf{Z}_L\end{bmatrix}\begin{bmatrix}\mathbf{u}_H\\ \mathbf{u}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{\mathbf{m}_H} & 0\\ 0 & \mathbf{Z}_{\mathbf{m}_L}\end{bmatrix}\begin{bmatrix}\mathbf{m}_H\\ \mathbf{m}_L\end{bmatrix} + \begin{bmatrix}\mathbf{e}_H\\ \mathbf{e}_L\end{bmatrix},$$

$$\text{with } \mathrm{cov}(u,m)=0 \quad (B_3)$$

$$\begin{bmatrix}\mathbf{y}_H\\ \mathbf{y}_L\end{bmatrix} = \begin{bmatrix}\mathbf{X}_H & 0\\ 0 & \mathbf{X}_L\end{bmatrix}\begin{bmatrix}\mathbf{b}_H\\ \mathbf{b}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_H & 0\\ 0 & \mathbf{Z}_L\end{bmatrix}\begin{bmatrix}\mathbf{u}_H\\ \mathbf{u}_L\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{\mathbf{m}_H} & 0\\ 0 & \mathbf{Z}_{\mathbf{m}_L}\end{bmatrix}\begin{bmatrix}\mathbf{m}_H\\ \mathbf{m}_L\end{bmatrix} + \begin{bmatrix}\mathbf{e}_H\\ \mathbf{e}_L\end{bmatrix},$$

$$\text{with } \mathrm{cov}(u,m) = \sigma_{um}\mathbf{A} \quad (B_4),$$

where subscript H (L) pertains to High (Low) hygiene conditions; $\mathbf{y}_{H(L)} = n_{H(L)}\times 1$ vector of observations ($n_{H(L)}$= number of records in H {L} = 203,323 {35,595}), $\mathbf{b}_{H(L)}$= p×1 vector of fixed effects (p = number of fixed effects classes in H {L} = 374 {198}), $\mathbf{u}_{H(L)}$ = q×1 vector of direct additive genetic effects (q = number of additive effects = 241,010), $\mathbf{m}_{H(L)}$ = d×1 vector of maternal genetic effects (d = total number of females =

123,452), $\mathbf{c}_{H(L)}$ = k×1 vector of maternal environmental effects (k=number of dams with offspring= 7,870), $\mathbf{e}_{H(L)}$ = $n_{H(L)}$×1 vector of residuals; $\mathbf{X}_{H(L)}$, $\mathbf{Z}_{H(L)}$, $\mathbf{Z}_{mH(L)}$ and $\mathbf{Z}_{cH(L)}$ denote the incidence matrices relating the observations to the corresponding fixed and random effects; **A** the additive relationship matrix. The vector of direct and maternal genetic effects was assumed to follow the multivariate normal distribution:

$$\left[\mathbf{u_H}\ \ \mathbf{u_L}\ \ \mathbf{m_H}\ \ \mathbf{m_L}\right]^T \sim N\left(\mathbf{0}_{2q+2d}, \mathbf{G} \otimes \mathbf{A}\right),$$

where $\mathbf{0}_N$ denotes a N×1 vector of 0s, $\otimes$ denotes the Kronecker product, A is the additive relationship matrix,

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{u_H} & \sigma_{u_H u_L} & \sigma_{u_H m_H} & \sigma_{u_H m_L} \\ \sigma_{u_H m_L} & \sigma^2_{u_L} & \sigma_{u_L m_H} & \sigma_{u_L m_L} \\ \sigma_{u_H m_H} & \sigma_{u_L m_H} & \sigma^2_{m_H} & \sigma_{m_H m_L} \\ \sigma_{u_H m_L} & \sigma_{u_L m_L} & \sigma_{m_H m_L} & \sigma^2_{m_L} \end{bmatrix}$$

is the 4×4 (co)variance matrix of direct and maternal genetic effects, $\sigma^2_{u_{H(L)}}$ denotes the direct genetic variance, $\sigma_{u_H u_L}$ the direct genetic covariance, $\sigma^2_{m_{H(L)}}$ the maternal genetic variance, $\sigma_{m_H m_L}$ the maternal genetic covariance and $\sigma_{u_{H(L)} m_{H(L)}}$ the direct-maternal genetic covariance. Maternal environmental effects for the two traits were assumed to follow the multivariate normal distribution: $\left[\mathbf{c_H}\ \ \mathbf{c_L}\right]^T \sim N\left(\mathbf{0}_{2k}, \mathbf{C} \otimes \mathbf{I_k}\right)$, where I_k is an identity matrix of order k,

$$\mathbf{C} = \begin{bmatrix} \sigma^2_{c_H} & \sigma_{c_H c_L} \\ \sigma_{c_H c_L} & \sigma^2_{c_L} \end{bmatrix}$$

is the 2×2 (co)variance matrix between maternal environmental effects, $\sigma^2_{c_{H(L)}}$ the maternal environmental variance and $\sigma_{c_H c_L}$ the maternal environmental covariance.

Finally, residuals, regarding different animals (males-females), were assumed independent between the two traits: $\left[\mathbf{e_H}\ \ \mathbf{e_L}\right]^T \sim N\left(\mathbf{0}_{2n}, \mathbf{R} \otimes \mathbf{I_n}\right)$, where $\mathbf{I_n}$ is an identity matrix of order n and

$$\mathbf{R} = \begin{bmatrix} \sigma^2_{e_H} & 0 \\ 0 & \sigma^2_{e_L} \end{bmatrix}$$

is the 2×2 residual (co)variance matrix, where $\sigma^2_{e_{H(L)}}$ is the residual variance.

5.3.2.2 Quadravariate analysis

Following the previous description, four animal models (Q1-Q4) were considered, for male and female body weight in H and L environments. In matrix notation these models can be described as follows:

$$\begin{bmatrix}\mathbf{y}_{H_1}\\ \mathbf{y}_{H_2}\\ \mathbf{y}_{L_1}\\ \mathbf{y}_{L_2}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{X}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{X}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{X}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{H_1}\\ \mathbf{b}_{H_2}\\ \mathbf{b}_{L_1}\\ \mathbf{b}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{Z}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{Z}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{Z}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{u}_{H_1}\\ \mathbf{u}_{H_2}\\ \mathbf{u}_{L_1}\\ \mathbf{u}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{H_1}\\ \mathbf{e}_{H_2}\\ \mathbf{e}_{L_1}\\ \mathbf{e}_{L_2}\end{bmatrix} \quad (Q_1)$$

$$\begin{bmatrix}\mathbf{y}_{H_1}\\ \mathbf{y}_{H_2}\\ \mathbf{y}_{L_1}\\ \mathbf{y}_{L_2}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{X}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{X}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{X}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{H_1}\\ \mathbf{b}_{H_2}\\ \mathbf{b}_{L_1}\\ \mathbf{b}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{Z}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{Z}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{Z}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{u}_{H_1}\\ \mathbf{u}_{H_2}\\ \mathbf{u}_{L_1}\\ \mathbf{u}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{c_H} & 0\\ 0 & \mathbf{Z}_{c_L}\end{bmatrix}\begin{bmatrix}\mathbf{c}_{H}\\ \mathbf{c}_{L}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{H_1}\\ \mathbf{e}_{H_2}\\ \mathbf{e}_{L_1}\\ \mathbf{e}_{L_2}\end{bmatrix} \quad (Q_2)$$

$$\begin{bmatrix}\mathbf{y}_{H_1}\\ \mathbf{y}_{H_2}\\ \mathbf{y}_{L_1}\\ \mathbf{y}_{L_2}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{X}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{X}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{X}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{H_1}\\ \mathbf{b}_{H_2}\\ \mathbf{b}_{L_1}\\ \mathbf{b}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{Z}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{Z}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{Z}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{u}_{H_1}\\ \mathbf{u}_{H_2}\\ \mathbf{u}_{L_1}\\ \mathbf{u}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m_H} & 0\\ 0 & \mathbf{Z}_{m_L}\end{bmatrix}\begin{bmatrix}\mathbf{m}_{H}\\ \mathbf{m}_{L}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{H_1}\\ \mathbf{e}_{H_2}\\ \mathbf{e}_{L_1}\\ \mathbf{e}_{L_2}\end{bmatrix},$$

with cov(u,m)=0 (Q_3)

$$\begin{bmatrix}\mathbf{y}_{H_1}\\ \mathbf{y}_{H_2}\\ \mathbf{y}_{L_1}\\ \mathbf{y}_{L_2}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{X}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{X}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{X}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{H_1}\\ \mathbf{b}_{H_2}\\ \mathbf{b}_{L_1}\\ \mathbf{b}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{H_1} & 0 & 0 & 0\\ 0 & \mathbf{Z}_{H_2} & 0 & 0\\ 0 & 0 & \mathbf{Z}_{L_1} & 0\\ 0 & 0 & 0 & \mathbf{Z}_{L_2}\end{bmatrix}\begin{bmatrix}\mathbf{u}_{H_1}\\ \mathbf{u}_{H_2}\\ \mathbf{u}_{L_1}\\ \mathbf{u}_{L_2}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m_H} & 0\\ 0 & \mathbf{Z}_{m_L}\end{bmatrix}\begin{bmatrix}\mathbf{m}_{H}\\ \mathbf{m}_{L}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{H_1}\\ \mathbf{e}_{H_2}\\ \mathbf{e}_{L_1}\\ \mathbf{e}_{L_2}\end{bmatrix},$$

with $\text{cov}(u,m) = \sigma_{um}\mathbf{A}$ (Q_4),

where subscript H (L) pertains to high (low) hygiene conditions and subscript 1 (2) pertains to male (female) BW; $\mathbf{y}_{H1(2)}$=$n_{H1(2)}$×1 vector of observations in H environment ($n_{H1(2)}$= number of male {female} records in H environment =99,330 {103,993}; $\mathbf{y}_{(L)1(2)}$ = $n_{L1(2)}$×1 vector of observations in L environment ($n_{L1(2)}$ = number of male {female} records in L environment =17,990 {17,605}); $\mathbf{b}_{H(L)}$ = $p_{H(L)}$×1 vector of fixed effects (p = number of fixed effects classes in H {L} environment = 374 {198}), $\mathbf{u}_{H(L)}$ = q×1 vector of direct additive genetic effects (q = number of additive effects = 241,010), $\mathbf{m}_{H(L)}$ = d×1 vector of maternal genetic effects (**d** = total number of females = 123,452), $\mathbf{c}_{H(L)}$= $k_{H(L)}$×1 vector of maternal environmental effects ($k_{H(L)}$ = number of dams with offspring in H {L} environment = 7,870 {3,822}), $\mathbf{e}_{H(L)}$ = $n_{H(L)}$×1 vector of residuals; $\mathbf{X}_{H(L)}$, $\mathbf{Z}_{H(L)}$, $\mathbf{Z}_{mH(L)}$ and $\mathbf{Z}_{cH(L)}$ denote the incidence matrices relating the observations to the corresponding fixed and random effects. The vector of direct and maternal genetic effects was assumed to follow the multivariate normal distribution: $\begin{bmatrix}\mathbf{u}_{H_1} & \mathbf{u}_{H_2} & \mathbf{u}_{L_1} & \mathbf{u}_{L_2} & \mathbf{m}_{H_1} & \mathbf{m}_{H_2} & \mathbf{m}_{L_1} & \mathbf{m}_{L_2}\end{bmatrix}^T \sim N\left(\mathbf{0}_{4q+4d}, \mathbf{G} \otimes \mathbf{A}\right)$, where $\mathbf{0}_N$ denotes a N×1 vector of 0s, $\otimes$ denotes the Kronecker product, **A** is the additive relationship matrix,

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{u_{H1}} & \sigma_{u_{H1}u_{H2}} & \sigma_{u_{H1}u_{L1}} & \sigma_{u_{H1}u_{L2}} & \sigma_{u_{H1}m_{H1}} & \sigma_{u_{H1}m_{H2}} & \sigma_{u_{H1}m_{L1}} & \sigma_{u_{H1}m_{L2}} \\ \sigma_{u_{H1}u_{H2}} & \sigma^2_{u_{H2}} & \sigma_{u_{H2}u_{L1}} & \sigma_{u_{H2}u_{L2}} & \sigma_{u_{H2}m_{H1}} & \sigma_{u_{H2}m_{H2}} & \sigma_{u_{H2}m_{L1}} & \sigma_{u_{H2}m_{L2}} \\ \sigma_{u_{H1}u_{L1}} & \sigma_{u_{H2}u_{L1}} & \sigma^2_{u_{L1}} & \sigma_{u_{L1}u_{L2}} & \sigma_{u_{L1}m_{H1}} & \sigma_{u_{L1}m_{H2}} & \sigma_{u_{L1}m_{L1}} & \sigma_{u_{L1}m_{L2}} \\ \sigma_{u_{H1}u_{L2}} & \sigma_{u_{H2}u_{L2}} & \sigma_{u_{L1}u_{L2}} & \sigma^2_{u_{L2}} & \sigma_{u_{L2}m_{H1}} & \sigma_{u_{L2}m_{H2}} & \sigma_{u_{L2}m_{L1}} & \sigma_{u_{L2}m_{L2}} \\ \sigma_{u_{H1}m_{H1}} & \sigma_{u_{H2}m_{H1}} & \sigma_{u_{L1}m_{H1}} & \sigma_{u_{L2}m_{H1}} & \sigma^2_{m_{H1}} & \sigma_{m_{H1}m_{H2}} & \sigma_{m_{H1}m_{L1}} & \sigma_{m_{H1}m_{L2}} \\ \sigma_{u_{H1}m_{H2}} & \sigma_{u_{H2}m_{H2}} & \sigma_{u_{L1}m_{H2}} & \sigma_{u_{L2}m_{H2}} & \sigma_{m_{H1}m_{H2}} & \sigma^2_{m_{H2}} & \sigma_{m_{H2}m_{L1}} & \sigma_{m_{H2}m_{L2}} \\ \sigma_{u_{H1}m_{L1}} & \sigma_{u_{H2}m_{L1}} & \sigma_{u_{L1}m_{L1}} & \sigma_{u_{L2}m_{L1}} & \sigma_{m_{H1}m_{L1}} & \sigma_{m_{H2}m_{L1}} & \sigma^2_{m_{L1}} & \sigma_{m_{L1}m_{L1}} \\ \sigma_{u_{H1}m_{L2}} & \sigma_{u_{H2}m_{L2}} & \sigma_{u_{L1}m_{L2}} & \sigma_{u_{L2}m_{L2}} & \sigma_{m_{H1}m_{L2}} & \sigma_{m_{H2}m_{L2}} & \sigma_{m_{L1}m_{L1}} & \sigma^2_{m_{L2}} \end{bmatrix}$$

is the 8×8 (co)variance matrix of direct and maternal genetic effects, $\sigma^2_{uH(L)_{1(2)}}$ denotes the direct genetic variance, $\sigma_{uH(L)_{1(2)}uL(H)_{2(1)}}$ the direct genetic covariance, $\sigma^2_{mH(L)_{1(2)}}$ the maternal genetic variance, $\sigma_{mH(L)_{1(2)}mL(H)_{2(1)}}$ the maternal genetic covariance and $\sigma_{uH(L)_{1(2)}mH(L)_{1(2)}}$ the direct-maternal genetic covariance. Maternal environmental effects for the two traits were assumed to follow the multivariate normal distribution:

$$\left[\mathbf{c}_{\mathbf{H}_1} \quad \mathbf{c}_{\mathbf{H}_2} \quad \mathbf{c}_{\mathbf{L}_1} \quad \mathbf{c}_{\mathbf{L}_2}\right]^T \sim N(\mathbf{0}_{4k}, \mathbf{C} \otimes \mathbf{I}_{\mathbf{k}}),$$

where $\mathbf{I}_{\mathbf{k}}$ is an identity matrix of order k,

$$\mathbf{C} = \begin{bmatrix} \sigma^2_{c_{H1}} & \sigma_{c_{H1}c_{H2}} & \sigma_{c_{H1}c_{L1}} & \sigma_{c_{H1}c_{L2}} \\ \sigma_{c_{H1}c_{H2}} & \sigma^2_{c_{H2}} & \sigma_{c_{H2}u_{L1}} & \sigma_{c_{H2}u_{L2}} \\ \sigma_{c_{H1}c_{L1}} & \sigma_{c_{H2}u_{L1}} & \sigma^2_{c_{L1}} & \sigma_{c_{L1}u_{L2}} \\ \sigma_{c_{H1}c_{L2}} & \sigma_{c_{H2}u_{L2}} & \sigma_{c_{L1}u_{L2}} & \sigma^2_{c_{L2}} \end{bmatrix}$$

is the 4×4 (co)variance matrix between maternal environmental effects, $\sigma^2_{cH(L)_{1(2)}}$ the maternal environmental variance and $\sigma_{cH(L)_{1(2)}cL(H)_{2(1)}}$ the maternal environmental covariance. Finally, residuals, regarding different animals (males-females), were assumed independent between the two traits: $\left[\mathbf{e}_{\mathbf{H}_1} \quad \mathbf{e}_{\mathbf{H}_2} \quad \mathbf{e}_{\mathbf{L}_1} \quad \mathbf{e}_{\mathbf{L}_2}\right]^T \sim N(\mathbf{0}_{4n}, \mathbf{R} \otimes \mathbf{I}_{\mathbf{n}})$, where $\mathbf{I}_{\mathbf{n}}$ is an identity matrix of order n and

$$\mathbf{R} = \begin{bmatrix} \sigma^2_{e_{H1}} & 0 & 0 & 0 \\ 0 & \sigma^2_{e_{H2}} & 0 & 0 \\ 0 & 0 & \sigma^2_{e_{L1}} & 0 \\ 0 & 0 & 0 & \sigma^2_{e_{L2}} \end{bmatrix}$$

is the 4×4 residual (co)variance matrix, where $\sigma^2_{eH(L)_{1(2)}}$ is the residual variance.

All analyses were carried out by the ASREML software (Gilmour et al, 2009). Estimates of direct (h^2) and maternal (h^2_m) heritability as well as maternal environmental effects (c^2) were calculated for each trait as ratios of estimates of direct additive (σ^2_u), maternal genetic (σ^2_m) and maternal environmental (σ^2_c) variances,

respectively to phenotypic variance (σ^2_p). The phenotypic variance accounts for the sum of all variance components, according to the model. The direct-maternal genetic correlation (r_{um}) was computed as the ratio of the estimate of direct-maternal genetic covariance (σ_{um}) to the product of the square roots of estimates of σ^2_u and σ^2_m. In addition, we follow Willham (1972) in calculating the total heritability (H^2_T) for BW as:

$$H^2_T = \frac{\sigma^2_u + 0.5\sigma^2_m + 1.5\sigma_{um}}{\sigma^2_p}.$$

5.3.2.3 Model evaluation criteria

Model comparison was carried out via three model evaluation criteria: the Akaike Information Criterion (AIC; Akaike, 1973), the Bayesian Information Criterion (BIC; Schwarz, 1978) and the conditional Akaike Information Criterion (cAIC; Vaida and Blanchard, 2005). The AIC gives an unbiased estimator of the Kullback-Leibler divergence of the current model from the true model. In addition, it can be shown that AIC chooses the model with the best short-term predictive ability (Stone, 1977). When a true model exists and lies within the range of models entertained in the analysis, the BIC (Schwarz, 1978) is consistent, in the sense that the true model will be selected as more data accrue. All model evaluation criteria are based upon the computation of the deviance (D): $D=-2\log(p(y|\hat{\theta}))=-2\log L$, where θ denotes the *px1* vector of the model parameters and $p(y|\hat{\theta})$ the likelihood of the data y evaluated at the maximum likelihood estimate $\hat{\theta}$. Akaike (1973) showed that the correct term for penalizing the deviance is twice the number of the model parameters *p*. Thus, he defined $AIC=-2\log L_i+2p$ as the model evaluation criterion. A Bayesian argument was utilized by Schwarz (1978) to prove that the appropriate penalization term is *p*log (*n*) thus defining: $BIC=-2\log L_i+p\log n$, where n is the number of data observations.

However, the determination of the number of the model parameters is non-trivial when random effects are of interest and are being estimated using methods such as BLUP. For such cases the AIC is shown in Crainiceanu and Ruppert (2004) to be asymptotically biased. In addition, Greven and Kneib (2010) showed that in linear mixed models AIC is a biased estimator of the Akaike information due to the non-open parameter space and the lack of independence between observations. The cAIC defined by Vaida and Blanchard (2005) as $cAIC=-2\log L_i+2\rho$ is asymptotically unbiased. Notice that ρ, the effective degrees of freedom (Hodges and Sargent, 2001), is given by the trace of the hat-matrix **H** which in the case of a simple additive animal model can be expressed as:

$$\rho = tr(\mathbf{H}) = tr\left(\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z+\lambda A^{-1} \end{bmatrix}^{-1}\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix}\right),$$

where $\lambda=\sigma^2_e/\sigma^2_u$. Minimizing the above criteria over a set of possible models can be seen as minimizing the average distance of an approximating model to the underlying truth (Greven and Kneib, 2010). Thus, the model with the smallest cAIC value is to be preferred.

5.4 Results

5.4.1 Bivariate analyses

Table 5.2 summarizes the estimated variance components and genetic parameters of both traits, along with the likelihoods and the various model evaluation criteria (AIC, BIC and cAIC) for the four bivariate models. Regarding the H environment, direct heritability ranged from 0.14 to 0.34, while maternal heritability was between 0.06 and 0.11, depending on the model. Furthermore, maternal environmental variance accounted for 0.05 of the total phenotypic variance of the trait. With regard to L environment, direct heritability varied from 0.27 to 0.46, while maternal heritability ranged from 0.03 to 0.06, depending on the model. Here, maternal environmental variance also accounted for 0.05 of the total phenotypic variance. A negative direct-maternal genetic correlation was detected in the H environment (-0.51) that was significantly higher ($P<0.05$) than in the L environment (-0.27). According to all model evaluation criteria, the best model included direct genetic effects and maternal environmental effects (model B_2). In this model, direct heritability was estimated as high as 0.14 and 0.27 in H and L, respectively. In both environments, maternal environmental variance accounted for 0.05 of the total phenotypic variance.

Genetic and phenotypic correlations between the two environments are presented in Table 5.3. Under a purely additive animal model, the additive genetic and phenotypic correlations were estimated as high as 0.48 and 0.19, respectively. The additive genetic correlation varied considerably between models with values ranging from 0.34 to 0.48, while the phenotypic correlation ranged from 0.10 to 0.19. The correlations between the maternal genetic effects ranged slightly from 0.77 to 0.79. According to model B_2, which was proposed by all evaluation criteria, the additive genetic correlation was estimated 0.35, indicating the presence of strong GEI. The respective phenotypic estimate was much lower (0.10), as the residual correlation between the two environments was set to zero, because residuals regarded different animals.

5.4.2 Quadravariate analysis

Table 5.4 summarizes the estimated variance components and genetic parameters of the four traits, along with the log-likelihoods and the model evaluation criteria (AIC, BIC and cAIC) for the three quadravariate models of analysis that converged. Direct heritability ranged from 0.14 to 0.40 for $BW_{H♂}$, while it was slightly higher for $BW_{H♀}$ ranging from 0.19 to 0.43. Given the standard errors, however, no sex differences were ascertained ($P>0.05$). Furthermore, results were consistent with those obtained in the bivariate analysis, with direct heritability being always higher in L than in H environment. Maternal heritability was estimated as high as 0.10 for $BW_{H♂}$ and 0.09 for $BW_{H♀}$, while it was 0.04 and 0.02 for $BW_{L♂}$ and $BW_{L♀}$, respectively. During model Q_2 that best fitted our data, direct heritability was estimated to be 0.14 and 0.19 for $BW_{H♂}$ and $BW_{H♀}$, respectively. Here, maternal environmental variance accounted for 0.08 and 0.07 of the total phenotypic variance of $BW_{H♂}$ and $BW_{H♀}$, respectively. In addition, no differences were detected in the maternal environmental variance between the two sexes in L environment ($c^2=0.06$ for both $BW_{L♂}$ and $BW_{L♀}$).

The genetic and phenotypic correlations are given in Table 5.5. The additive genetic correlations (r_u) between the sexes in H environment were high, ranging from 0.81 to 0.92, and this was also the case for the L environment. Likewise, the estimates of the maternal genetic correlations between $BW_{H♂}$ and $BW_{H♀}$ (0.97) and between $BW_{L♂}$ and $BW_{L♀}$ (0.92) were also high. Notably all the additive genetic correlations

across sexes, between the environments were low ranging from 0.24 ($BW_{H♂}$-$BW_{L♀}$) to 0.50 ($BW_{H♂}$-$BW_{L♂}$). As in the bivariate case, the residual correlation between the four traits was set to zero thus leading to low phenotypic correlations across the traits. Under model Q_2, the direct genetic correlation between sexes in the same environment were 0.81 and the correlations between the maternal environmental effects were higher (0.94 and 0.87 between $BW_{H♂}$-$BW_{H♀}$ and $BW_{L♂}$-$BW_{L♀}$, respectively). Conversely, the estimates of the same sex in the two environments were significantly smaller (e.g. 0.45 for $BW_{H♂}$-$BW_{L♂}$ and 0.31 for $BW_{H♀}$-BW_{L}).

5.5 Discussion

This appears to be the first study reporting significant GEI for an important economic trait such as body weight arising from different hygiene conditions. Results of the present study have shown animals' average performance i.e. body weight to be severely affected by the hygiene level conditions, with chicken raised in H being significant heavier than their L counter mates. Apart from the mean trait, performance was more variable in the L vs. the H environment, manifested in terms of phenotypic, additive genetic variance and direct heritability estimates. Raising broilers under low hygiene level conditions may pose a significantly stressful environment. Environments of this kind are reported to increase both the phenotypic and the genetic variation in a number of quantitative characters (Barker and Krebs, 1995; Zhivotovsky et al, 1996; Sgro and Hoffman, 1998), a finding that has both evolutionary and breeding implications. From an evolutionary point of view, increased genetic variation translates to increased adaptive potential and thereby survival probability in adverse environments (Imascheva et al., 1998). Following the old breeder's dogma that "the character required is best selected for under environmental conditions which favor its fullest expression" i.e. exhibiting highest heritability, selection should mainly practiced in the L environment. This may be valid, however, only when the genetic correlation between the two environments is high ($r_G \geq 0.8$) i.e. when the GEI is of no or very little importance. Given the low ($r_G \sim 0.40$) genetic correlation across the two environments estimated herein, there seems to be significant GEI for the trait that should be appropriately accounted for during estimation of breeding values.

Although first demonstrated for hygiene conditions here, GEI is repeatedly reported as being of importance in many aspects (or variety) of environmental conditions and traits, in broilers. In agreement with our results, Banos et al. (2006) have reported higher heritability estimates of BW_{35} in commercial than in selection hygiene environments (0.32 vs. 0.22). In a cross between two genetically different broiler dam lines, raised in cold and normal temperatures, Pakdel et al. (2005) obtained higher heritability estimates in the normal than in the cold conditions (0.50 vs. 0.42). Notably, in accordance with our results, most of the other traits displayed higher direct heritabilities in the harsh (cold) than in the normal environment. In addition, Pakdel et al. (2005) have also estimated a low (0.29) genetic correlation for BW in broilers between cold and normal temperature conditions. In contrast to our findings, Zerehdaran et al. (2005) reported genetic correlations ranging from 0.78 to 0.89 for BW at three different ages and two distinct housing systems. High genetic correlations (0.74 to 0.98) are also reported by N'Dri et al. (2007) for BW at 8 weeks of age in slow-growing chicken between three environments ranged from. The controversial literature data together with our findings suggest that the effect of harsh conditions on genetic variability might be trait-specific and depends on the kinds and levels of the extreme environmental conditions. Thus, generalizations about the behavior of genetic variation under extreme conditions should be avoided. Differences in estimates throughout the

studies may reflect differences in genotype(s) used (crosses, fast or slow growing lines) as well as environmental conditions (housing system, temperature).

Finally, we have tested the hypothesis that GEI might be sex-specific. To do so, we have applied a number of animal models and a quadravariate approach. Here, body weight per sex and environment was treated as distinct traits assuming no residual correlation between. During this approach, direct genetic correlations between sexes both within as well as across the two environments were as minimum as 0.81. Furthermore, no difference between sex estimates, whether heritabilities or genetic correlations, were established. All the above imply that GEI is not sex related and thus no sex specific selection strategies should be envisaged. This finding was further supported by analysis of variance (ANOVA) where sex (S), environment (E) and the sex by environment (SxE) term were treated as fixed effects. During ANOVA, the SxE term explained only 2.7% of the trait variance (results not shown).

In conclusion, a strong GEI interaction for BW due to hygiene conditions was revealed suggesting that there is sufficient phenotypic plasticity for the trait under a range of environmental conditions (when the two hygiene levels are considered as two extremes). Obtaining records in both environments represents an important information source regardless of the breeding goal i.e. breeding for the high (selection nucleus), low (commercial) or a range of hygiene conditions falling within the two extremes. When omitted, the loss of genetic gain due to GEI is expected to be of appreciable magnitude. Simulation studies (Mulder and Bijma, 2005) have shown that progeny-testing rather than sib-testing schemes are preferable when there are low to moderate heritabilities ($h^2 \leq 0.30$), relatively short generation intervals of progeny-tested sires (≤ 1.7) and moderate to severe GEI ($r_u \leq 0.80$). It has been further shown (Mulder and Bijma, 2005) that increasing the number of progeny per sire, the proportion of selected sires and the population size in the harsh environment, minimizes the loss in genetic gain due to GEI.

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, Edition Petrov and Csaki, Budapest, Hungary.

Banos, G., Avendano, S. and Olori, V. 2006. Time dependent genetic parameters for broiler chicken body weight measured in selection and commercial environments. *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*, *Belo Horizonte, Brazil.*

Barker, J.S.F. and Krebs, R.A. 1995. Genetic variation and plasticity of thorax length and wing length in *Drosophila aldrichi* and *Drosophila Buzzattii. Journal of Evolutionary Biology* 8: 689-709.

Cahaner, A. and Leenstra, F.R. 1992. Effects of high temperature on growth and efficiency of male and female broilers from lines selected for weight gain, favorable feed conversion and high or low fat content. *Poultry Science* 71: 1237-1250.

Case, L.A., Kelly, M.J., Miller, S.P. and Wood, B.J. 2010. Genotype x Environment interaction as it relates to egg production in turkeys. *Journal of Animal Science* 88: 1957- 1966.

Crainiceanu, C. and Ruppert, D. 2004. Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society, Series B* 66: 165–85.

Falconer, D.S. and Mackay, T.F. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group, Essex, UK.

Fordyce, J.A. 2006. The evolutionary consequences of ecological interactions mediated through phenotypic plasticity. *Journal of Experimental Biology* 209: 2377–2383.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. 2009. ASReml User Guide Release 3.0. *VSN International Ltd*, Hemel Hempstead, UK.

Greven, S. and Kneib, T. 2010. On the behaviour of marginal and conditional Akaike Information Criteria in Linear Mixed Models. *Biometrika* 97: 773-789.

Havenstein, G.G., Ferket, P.R., Scheideler, S.E. and Larson, B.T. 1994. Growth, livability and feed conversion of 1957 vs 1991 broilers when fed "typical" 1957 and 1991 broiler diets. *Poultry Science* 73: 1795-1804.

Hodges, J.S. and Sargent, D.J. 2001. Counting degrees of freedom in hierarchical and other richly-parameterized models. *Biometrika* 88: 367-379.

Imasheva, A., Loeschcke, V., Zhivotovsky, L and Lazebny, O. 1998. Stress temperatures and quantitative variation in *Drosohila melanogaster*. *Heredity* 81: 246-253.

Leenstra, F.R. and Cahaner, A. 1992. Effects of low, normal and high temperatures on slaughter yield of broilers from lines selected for high weight gain, favorable feed conversion and high or low fat content. *Poultry Science* 71: 199-206.

Long, N., Gianola, D., Rosa, G.J.M., Weigel, K.A. and Avendano, S. 2008. Marker-assisted assessment of genotype by environment interaction: A case study of single nucleotide polymorphism-mortality association in broilers in two hygiene environments. *Journal Animal Science* 86: 3358-3366.

Mulder, H.A. and Bijma, P. 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. *Journal Animal Science* 83: 49-61.

Mulder, H.A., Veerkamp, R.F., Ducro, B.J., van Arendonk, J.A.M. and Bijma, P. 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. *Journal of Dairy Science* 89: 1740-1752.

N'Dri, A.L., Sellier, N., Tixier-Boichard, M., Beaumont, C and Mignon-Grasteau, S. 2007. Genotype by environment interactions in relation to growth traits in slow growing chickens. *Genetics Selection Evolution* 39: 513-528.

Pakdel, A., van Arendonk, J.A.M., Vereijken, A.L.J. and Bovenhuis, H. 2005. Genetic parameters of ascites-related traits in broilers: effect of cold and normal temperature conditions. *British Poultry Science* 46: 35-42.

Robertson, A. 1959. The sampling variance of the genetic correlation coefficient. *Biometrics* 15: 469-485.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464

Settar, P., Yalcin, S., Turkmut, L., Ozkan, S and Cahanar, A. 1999. Season by genotype interaction related to broiler growth rate and heat tolerance. *Poultry Science* 78: 1353-1358.

Sgro, C.M. and Hoffmann, A.A. 1998. Effects of temperature extremes on genetic variances for life history traits in *Drosohila melanogaster* as determined from parent-offspring comparisons. *Journal of Evolutionary Biology* 11: 1-20.

Stone, M. 1977. An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal of the Royal Statistical Society, Series B* 30(1): 44-47.

Vaida, F. and Blanchard, S. 2005. Conditional Akaike information for mixed-effects models. *Biometrika* 92: 351-370.

Willham, R.L. 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. *Journal Animal Science* 35: 1288-1293.

Yalcin, S., Testik, A., Ozkan, S, Settar, P., Celen, F. and Cahaner, A. 1997. Comparative evaluation of three commercial broiler stocks in hot temperate climates. *Poultry Science* 76: 921-929.

Ye, X., Avendano, S., Dekkers, J.C.M. and Lamond, S.J. 2006. Association of twelve immune-related genes with performance of three broiler lines in two different hygiene environments. *Poultry Science* 85: 1555-1569.

Zerehdaran, A.L., Vereijken, A.L.J., van Arendonk, J.A.M. and van der Waaij, E.H. 2005. Effect of age and housing system on genetic parameters for broiler carcass traits. *Poultry Science* 84: 833-838.

Zhivotovsky, L., Feldman, M.W. and Bergman, A. 1996. On the evolution of phenotypic plasticity in a spatially heterogeneous environment. *Evolution* 50: 547-558.

Tables and Figures

Table 5.1 Description of the data set in a high (H) and a low (L) hygiene environment and the pedigree used.

Environment	H	L
No. of total animals	205,415	40,082
No. of animals with records	203,323	35,595
No. of males with records	99,330	17,990
No. of females with records	103,993	17,605
No. of sires	980	665†
No. of dams	7,870	3,822†

† common to both environments

Table 5.2 Estimates of variance components, genetic parameters, log-likelihoods and model selection criteria, under four bivariate models for the body weight (BW) of broiler chicken at 35 days of age in a high (H) and a low (L) hygiene environment.

Model	Env	σ^2_u	σ^2_m	σ_{um}	σ^2_c	σ^2_{fs}	σ^2_e	σ^2_p	h^2	h^2_m	$\frac{\lvert\sigma_{um}\rvert}{\sigma^2_p}$	r_{um}	c^2	$\frac{\sigma^2_{fs}}{\sigma^2_p}$	H^2_T	logL	AIC	BIC	cAIC	ρ
	H	23,766	-	-	-	-	46,855	70,621	0.34	-	-	-	-	-	0.34	-1,438,690	2,877,382	2,877,392	2,932,999	27,809
B_1		(600)					(400)	(400)	(0.01)						(0.01)					
	L	40,661	-	-	-	-	47,231	87,893	0.46	-	-	-	-	-	0.46					
		(600)					(800)	(800)	(0.02)						(0.02)					
	H	9,493	-	-	3,020	-	53,452	65,965	0.14	-	-	-	0.05	-	0.14	**-1,438,317**	**2,876,638**	**2,876,659**	**2,904,981**	**14,172**
B_2		(500)			(200)		(300)	(300)	(0.01)				(0.01)		(0.01)					
	L	22,302	-	-	4,532	-	55,829	82,662	0.27	-	-	-	0.05	-	0.27					
		(1,100)			(500)		(800)	(800)	(0.02)				(0.01)		(0.02)					
	H	10,700	3,941	-	-	-	52,923	67,563	0.16	0.06	-	-	-	-	0.19	-1,438,382	2,876,768	2,876,789	2,909,536	16,384
B_3		(600)	(600)				(300)	(300)	(0.01)	(0.01)					(0.01)					
	L	31,871	2,450	-	-	-	51,504	85,825	0.37	0.03	-	-	-	-	0.39					
		(1,200)	(700)				(900)	(900)	(0.02)	(0.01)					(0.02)					
	H	13,636	7,375	-5,086	-	-	51,485	67,410	0.20	0.11	0.08	-0.51	-	-	0.14	-1,438,321	2,876,648	2,876,679	2,914,919	19,135
B_4		(900)	(600)	(500)			(300)	(300)	(0.01)	(0.01)	(0.01)	(0.03)			(0.01)					
	L	34,068	4,880	-3,516	-	-	50,421	85,852	0.40	0.06	0.04	-0.27	-	-	0.36					
		(1,400)	(900)	(900)			(900)	(900)	(0.02)	(0.01)	(0.01)	(0.06)			(0.02)					

σ^2_u: direct additive genetic variance; σ^2_m: maternal genetic variance; σ_{um} = direct-maternal genetic covariance; σ^2_c: maternal environmental variance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; h^2_m: maternal heritability; r_{um}: direct-maternal genetic correlation; c^2: maternal environmental variance as proportion of σ^2_p; H^2_T: total heritability; logL: natural log-likelihood; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; cAIC: conditional Akaike Information Criterion; ρ: effective degrees of freedom; standard errors in parenthesis;

Table 5.3 Phenotypic correlation coefficient (r_p), direct genetic correlation coefficient (r_u), maternal genetic correlation coefficient (r_m) and correlation coefficients between maternal environmental effects (r_c) for body weights (BW) of broilers at 35 days of age in a high and a low hygiene environment, under four bivariate models of analysis.

Models	r_u	r_m	r_c	r_{fs}	r_p
B_1	0.48 (0.02)	-	-	-	0.19 (0.01)
B_2	0.35 (0.05)	-	0.48 (0.05)	-	0.10 (0.01)
B_3	0.34 (0.05)	0.77 (0.09)	-	-	0.11 (0.01)
B_4	0.37 (0.05)	0.79 (0.08)	-	-	0.12 (0.01)

standard errors in parenthesis

Table 5.4 Estimates of variance components, genetic parameters, log-likelihoods and model selection criteria under three quadravariate models for the body weight (BW) of male (♂) and female (♀) broiler chicken at 35 days of age in a high (H) and a low (L) hygiene environment.

Model	Trait	σ^2_u	σ^2_m	σ^2_c	σ^2_{fs}	σ^2_e	σ^2_p	h^2	h^2_m	c^2	$\frac{\sigma^2_{fs}}{\sigma^2_p}$	H^2_T	logL	AIC	BIC	cAIC	ρ
Q_1	♂H	19,929 (600)	-	-	-	30,233 (300)	50,161 (300)	0.40 (0.01)	-	-	-	0.40 (0.01)	-1,394,618	2,789,244	2,789,261	2,818,924	14,842
	♀H	19,974 (600)	-	-	-	26,372 (300)	46,346 (300)	0.43 (0.01)	-	-	-	0.43 (0.01)					
	♂L	37,257 (1,100)	-	-	-	44,233 (900)	81,490 (900)	0.46 (0.02)	-	-	-	0.46 (0.02)					
	♀L	40,527 (1,200)	-	-	-	42,225 (950)	82,752 (950)	0.49 (0.02)	-	-	-	0.49 (0.02)					
Q_2	♂H	6,681 (500)	-	3,643 (200)	-	35,980 (300)	46,304 (300)	0.14 (0.01)	-	0.08 (0.01)	-	0.14 (0.01)	**-1,393,795**	**2,787,606**	**2,787,640**	**2,802,842**	**7,622**
	♀H	7,927 (500)	-	2,911 (150)	-	31,721 (300)	42,559 (300)	0.19 (0.01)	-	0.07 (0.01)	-	0.19 (0.01)					
	♂L	20,210 (1,100)	-	4,845 (600)	-	51,869 (900)	76,924 (900)	0.26 (0.02)	-	0.06 (0.01)	-	0.26 (0.02)					
	♀L	23,391 (1,100)	-	4,437 (600)	-	50,105 (900)	77,933 (900)	0.30 (0.02)	-	0.06 (0.01)	-	0.30 (0.02)					
Q_3	♂H	9,034 (600)	5,252 (200)	-	-	36,765 (300)	51,051 (300)	0.18 (0.01)	0.10 (0.01)	-	-	0.23 (0.01)	-1,393,900	2,787,816	2,787,850	2,806,213	9,203
	♀H	9,055 (600)	4,180 (200)	-	-	31,736 (300)	45,371 (300)	0.21 (0.01)	0.09 (0.01)	-	-	0.25 (0.01)					
	♂L	28,386 (1,200)	3,341 (700)	-	-	48,395 (900)	80,121 (900)	0.35 (0.03)	0.04 (0.01)	-	-	0.37 (0.03)					
	♀L	32,644 (1,300)	2,003 (700)	-	-	46,123 (900)	80,770 (900)	0.40 (0.03)	0.02 (0.01)	-	-	0.41 (0.03)					

σ^2_u: direct additive genetic variance; σ^2_m: maternal genetic variance; σ^2_c: maternal environmental variance; σ^2_e: residual variance; σ^2_p: phenotypic variance in g^2; h^2: direct heritability; h^2_m: maternal heritability; c^2: maternal environmental variance as proportion of σ^2_p; H^2_T: total heritability; logL: natural log-likelihood; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; cAIC: conditional Akaike Information Criterion; ρ: effective degrees of freedom; standard errors in parenthesis;

Table 5.5 Phenotypic correlation coefficient (r_p), direct genetic correlation coefficient (r_u), maternal genetic correlation coefficient (r_m) and correlation coefficients between maternal environmental effects (r_c) , for body weights of male (♂) and female (♀) broiler chicken at 35 days of age, in a high (H) and a low (L) hygiene environment, under three quadravariate models of analysis.

Models	Correlation	♂H, ♀H	♂L, ♀L	♂H, ♂L	♀H, ♀L	♂H, ♀L	♂L, ♀H
Q_1	r_u	0.92 (0.01)	0.92 (0.01)	0.50 (0.03)	0.43 (0.03)	0.41 (0.03)	0.43 (0.03)
Q_2		0.81 (0.02)	0.81 (0.02)	0.45 (0.05)	0.31 (0.05)	0.28 (0.06)	0.33 (0.05)
Q_3		0.85 (0.02)	0.92 (0.02)	0.39 (0.05)	0.25 (0.05)	0.24 (0.05)	0.30 (0.05)
Q_3	r_m	0.97 (0.01)	0.92 (0.06)	0.71 (0.06)	0.73 (0.08)	0.75 (0.09)	0.70 (0.07)
Q_2	r_c	0.94 (0.02)	0.87 (0.07)	0.42 (0.06)	0.40 (0.07)	0.47 (0.07)	0.45 (0.06)
Q_1	r_p	0.38 (0.01)	0.44 (0.02)	0.21 (0.01)	0.19 (0.01)	0.18 (0.01)	0.19 (0.01)
Q_2		0.20 (0.01)	0.30 (0.02)	0.12 (0.01)	0.10 (0.01)	0.10 (0.01)	0.10 (0.01)
Q_3		0.26 (0.01)	0.38 (0.02)	0.14 (0.01)	0.11 (0.01)	0.10 (0.01)	0.12 (0.01)

standard errors in parenthesis

6. Concluding remarks

6.1 Statistical methodology

The aim of quantitative genetics is to separate variances between animals into causal components due to different modes of action. Typical modes of action include additive, maternal (genetic or environmental) as well as non-additive genetic effects (e.g. dominance, epistatic) (Palucci et al., 2007). Extensions may also comprise of parent-of-origin i.e. imprinting effects (e.g. Essl and Voith, 2002), mitochondrial i.e. cytoplasmatic (e.g. Schutz et al., 1992) and sex-linked inheritance (Fairbairn and Roff, 2006). Apart from genetic, social effects may also be present for socially affected traits (Muir and Schinkel, 2002). In cases where non genetic relationships between direct and maternal effects are present, the so called maternal animal model (Quaas and Pollak, 1980; Henderson, 1988) has been proposed. Extensions here include correlation between maternal environmental effects of related dams (Quintanilla et al., 1999). In conclusion, there is a tendency towards using more complicated and sophisticated models in an effort to capture all possible sources of variation.

Whatever the mode of action, interactions between the various effects may also be of (some) importance. This interaction is quantified via appropriate statistical measures like the covariance. Existent theory presumes various assumptions regarding the (co)variance among the random effects. Real data however, may significantly violate assumptions of this kind. Apart from the usual covariance between the additive and maternal effects that seems to be successfully modeled, failure to account for existent covariances may affect parameter estimation during VCE and BLUP. For instance, chapter 2 has demonstrated that the common assumption of independence between the additive genetic effects and the maternal environmental effects does not always hold. Whatever the sources of such covariance(s), results imply that a more thorough examination of the data should be pursued, attempting to reveal possibly 'hidden' covariance structure(s). In our case, this was only feasible by employment of WinBUGS, a salient feature that should be appreciably acknowledged. Unfortunately, the inability to analyze large data sets limits the application of this software, reserving it an explanatory tool of latent structures on handy sub-sets of the original data. A promising alternative has emerged in the form of INLA, which certainly requires further development and testing. Regardless of the scholar (frequentist or Bayesian), there seems to be a need for more flexible algorithms and software(s) being capable to fit more sophisticated (co)variance structures. Although the Moore's law relates computation efficiency to hardware advances, the remarkable present computational efficiency should be mostly attributed to improvement(s) in algorithms.

Statistics has traditionally been the mentor of Animal Breeding and shall continue to do so. Advances attained there should be quickly and efficiently transferred to Animal Breeding and this can be accomplished only via qualified breeders with a good background in statistics. The prevailing view that molecular genetics would solve several problems in animal breeding has gradually lost ground (Eisen, 2008). With the advent of GS, the need of scientists combining knowledge of molecular genetics and statistical methods becomes even more imperative.

6.2 Biological aspects

Results obtained herein have proved body weight (BW) a trait with salient biological features:

a) A number of studies (e.g. Muir et al., 2008) have shown that commercial pure lines display significant absence of rare alleles i.e. lower genetic diversity compared to non commercial chickens. The levels of additive genetic variance estimated herein do not fully comply with such results since a moderate heritability for the trait was estimated. Despite the intense levels of selection, the closed populations and the industry consolidation, there are no significant signs of genetic erosion due to selection, at least in the population under study, ensuring future genetic progress. In contrast to what is generally believed, nucleotide polymorphism and genome analysis show that past artificial selection has not substantially reduced the genetic variation of the domestic chicken genome when compared to its wild ancestor, the red jungle fowl (International Chicken Polymorphism Map Consortium, 2004; Sawai et al., 2010). New mutations may provide the observed genetic variability and contribute to the lack of a perceived selection limit for growth and reproduction traits (Cahaner et al., 1996). Although no selection limits are in sight, side effects from selection for enhanced BW are already observed in the species. These include infertility, increased number of defective eggs, higher mortality, leg (e.g. tibial dyschondroplasia) and heart disorders and ascites (Rauw et al., 1998). To soften the impact, a more holistic approach is currently being practiced comprising of various selection criteria.

b) As in most livestock species, BW in broilers is under the genetic control of both additive and maternal genetic effects. The two effects are negatively correlated so that selection for increased phenotypic performance will be ablated by an amount that equals to the direct-genetic covariance (Willham, 1972). This negative correlation is the result of the pleiotropic action of genes (antagonistic pleiotropy; Roff, 2002). Its role is of tremendous importance for maintaining genetic variance and preventing the gradual shift towards higher weights.

c) BW is a trait with significant phenotypic plasticity and this was observed under environments which reflect differentiated rearing conditions with regard to hygiene. Although the two environments studied here do not represent natural habitats, rearing birds under extreme conditions always pose a real challenge for these organisms. Under challenging (low) hygiene conditions, birds exhibit higher phenotypic and most importantly genetic variation, in full accordance with their living free relatives.

All the above findings teach us a valuable lesson: apart from implications of artificial selection, animals at our disposal obey first and above all the rules of natural selection and evolution. A better understanding of these rules is necessary not only for exploring strategies for further improvement but also for recognizing the biological limits of our genetic material. Species and traits evolve and so does animal breeding. New developments in genomics have recently made possible the application of GS. The method has major advantages: it allows for early selection (reduction of the generation interval) and is independent on sex-linked expression and/or family structures. Most importantly, through the availability of high density SNP chips (Groenen et al. 2011) it has gained increased accuracy. The latter chips are used for parental genotyping which is then combined with the use of imputation of progeny genotypes obtained from low density chips in attempts to reduce the "phase" problem. Current statistical approaches are focused on additive genetic effects; dominance and epistatic effects require additional exploration. The focus is selection at younger ages and for specific traits such as feed efficiency, disease resistance, robustness and so forth, which are difficult to improve via traditional methods. The first results using 30-K SNP chips were obtained from a commercial line used for training, validation and selection. They have shown improved accuracy of prediction at a young age and therefore resulted in increased

genetic gain(s). The key challenge to adopting this methodology is the high density genotyping cost (around 100-250 Euros per bird; Avendano et al., 2010) but GS must prove its advantages over traditional methods including cost benefits (Preisinger, 2011).

Throughout this thesis it has become profound that increasingly complex models are eagerly tested as better approximations to reality. Bayesian and REML-based methods can certainly cooperate in this task unraveling existent data structures and providing effective estimates of the parameters in question. But how can all these models be effectively evaluated? Certainly, cAIC which was introduced in this thesis can be of great assistance. Towards this direction, further simulation studies should be conducted in order to investigate the impact of different pedigree structures as well as the existence of covariances between the random effects on the estimation of the effective degrees of freedom. Reaching an end and always questions and necessities arise. Can the antagonism between direct and maternal genetic effects generate novel variation having thus an important role towards the maintenance of variation as well as the evolution of broilers' body weight? In this thesis a trend of differential direct-maternal genetic correlation between the sexes was revealed for body weight of broilers at 35 days of age, while no differences between genders were detected at 7 days of age. Could this finding be an indication of a mechanism involved in the sexual dimorphism of chicken? Statistical analysis is always dependent on the available data structures. In order to have a clearer image, body weights measured at least at three different ages from 7 to 35 days should be appropriate to unveil via longitudinal models of analysis an aspect of this phenomenon. Genomic data could also be of assistance towards this direction to unlock the mechanisms involved. A second finding of this thesis was the existence of strong genotype by hygiene environment interaction for the respective trait of importance (BW). Although, climatic, hygiene or nutritional conditions have been reported in the established literature to be a source for GEI, almost never have all (or some of) these conditions been tested simultaneously. Thus, measurement of a combination of environmental conditions would provide a more appropriate approximation to real extreme environments of raising broilers and would assist in describing the pattern of phenotypic expression of genotypes across a range of environments through a norm of reaction. Nowadays, although the body weight of boilers remains the most important economical trait for the particular industry, selection of an increasingly large number of traits has become the aim of animal breeding schemes. Feed conversion rate, hen housed production, breast percentage in carcass are only a few traits that are in the scope. Thus, it would be of extreme interest a dataset regarding more traits and via multivariate analysis to estimate among others the genetic and phenotypic correlations between them.

REFERENCES

Avendano, S., Watson, K.A. and Kranis, A. 2010. Genomics in Poultry Breeding – From Utopias to Deliverables. *9th Congress on Genetics Applied to Livestock Production*, Leipzig, Germany.

Cahaner A, Smith E.J., Swenson S., Lamont S.J. 1996 Associations of individual genomic heterozygosity, estimated by molecular fingerprinting, and of dam major histocompatibility complex with growth and egg production traits in layer chickens. *Poultry Science* 75:1463–1467.

Eisen, E. 2008. Can we rescue an endangered species? *Journal of Animal Breeding and Genetics* 125: 1–2.

Essl, A. and Voith, K. 2002. Estimation of variance components due to imprinting effects with DFREML and VCE: results of a simulation study. *Journal of Animal Breeding and Genetics* 119: 1–14.

Fairbairn, D.J. and Roff, D.A. 2006. The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. *Heredity* 97: 319–328.

Groenen, M.A.M., Megens, H.-J., Zare, Y., Warren, W.C., Hillier, L.W., Crooijmans, R.P.M.A., Vereijken, A., Okimoto, R., Muir, W.M. and Cheng, H.H. 2011. The development and characterization of a 60K SNP chip for chicken. *BMC Genomics* 12: 274 doi:10.1186/1471-2164-12-274.

Henderson, C.R. 1988. Theoretical basis and computational methods for a number of different animal models. *Journal of Dairy Science* 71: 1-16.

International Chicken Polymorphism Map Consortium 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432: 717–722.

Muir, W.M. and Schinckel, A. 2002. Incorporation of competitive effects in breeding programs to improve productivity and animal well being. *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, August, 2002, Session 14: 1-4.

Muir, W.M., Wong, G.K.S., Zhang, Y., Wang, J., Groenen, M.A.M., Crooijmans, R.P.M.A., Megens, H.J., Zhang, H. Okimoto, R., Vereijken, A., Jungerius, A., Albers, G.A.A., Lawley, C.T., Delany, M.E., MacEachern, S. and Cheng, H.H. 2008. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proceedings of the National Academy of Sciences USA* 105: 17312-17317.

Palucci, V., Schaeffer, L.R., Miglior, F. and Osborne, V. 2007. Non-additive genetic effects for fertility traits in Canadian Holstein cattle. *Genetics Selection Evolution* 39: 181-193.

Preisinger, R. 2011. Genome-wide selection in poultry. *Animal Production Science* 52(3): 121-125.

Quaas, R.L. and Pollak, E. J. 1980. Mixed Model Methodology for Farm and Ranch Beef Cattle Testing Programs. *Journal of Animal Science* 51:1277-1287.

Quintanilla, R., Varona, L., Pujol, M.R. and Piedrafita, J. 1999. Maternal animal model with correlation between maternal environmental effects of related dams. *Journal of Animal Science* 77: 2904-2917.

Rauw, W.M., Kanis, E., Noordhuizen-Stassen, E.N. and Grommers, F.J. 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Production Science* 56: 15-33.

Roff, D.A. 2002 Life history evolution. Sinauer, Sunderland, MA.

Sawai, H., Kim, H.L., Kuno, K., Suzuki, S., Gotoh, H., Takada, M., Takahata, N., Satta, Y. and Akishinonomiya, F. 2010. The Origin and Genetic Variation of Domestic Chickens with Special Reference to Junglefowls Gallus g. gallus and G. varius. *PLoS ONE* 5(5): e10639. doi:10.1371/journal.pone.0010639.

Schutz, M.M., Freeman, A.E., Beitz, D.C. and Mayfield, J.E. 1992. The importance of Maternal Lineage on Milk Yield Traits of Dairy Cattle. *Journal of Dairy Science* 75: 1331-1341.

Willham, R.L. 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. *Journal of Animal Science* 35: 1288-1293.

APPENDIX

Codes for functions in ASREML, R and WinBUGS

Simulation for chapter 2 via R

```
tag<-c(1:2330) #individuals in pedigree
#parents given externally
read.table(file="sire.txt",header=TRUE)
read.table(file="dam.txt",header=TRUE)
read.table(file="c2.txt",header=TRUE) #connecting dams-offspring for the c2 effects
n<-2330   #total number of individuals in the pedigree
f<-90   #number of parents
d<-f+1
μ<-c(rep(10,2330)) #fixed effect
#random effects
s2u<-7
su<-sqrt(s2u)
s2c<-3
sc<-sqrt(s2c)
s2e<-32
se<-sqrt(s2e)
r<- -0.80
s2p<-s2u+s2c+s2e
h2<-s2u/s2p
c2<-s2c/s2p
N<-30  #number of iterations
#matrices for saving samples
u.hat<-matrix(0,N,n)
c.hat<-matrix(0,N,n)
e.hat<-matrix(0,N,n)
p.hat<-matrix(0,N,n)
s2u.hat<-matrix(0,N,1)
s2c.hat<-matrix(0,N,1)
s2e.hat<-matrix(0,N,1)
s2p.hat<-matrix(0,N,1)
h2.hat<-matrix(0,N,1)
c2.hat<-matrix(0,N,1)
start <- Sys.time() #init time
for (i in 1:N)
{#c2 effects
c1<-rnorm(f-20,0,1)*sc #distribution
c0<-c(rep(0,90)) #0 c2 for parents
c<-c(c0,c1[91:2330]) #joint c2 of parents and offspring
```

```
#breeding values simulation
u1<-rnorm(f,0,su) #breeding values of parents
mu<-(sqrt(0.5*s2u)/sc)*r*c #mean of the offspring bvs
ms<-rnorm((n-f),mu,sqrt((1-r^2)*0.5*s2u)) #mendelian sampling for offspring bvs
u2<-1/2*(u1[sire[d:n]]+u1[dam[d:n]])+ms #breeding values of offspring
u<-c(u1,u2) #breeding values joint
e<-rnorm(n,0,1)*se #residuals
y<-μ+u+c+e #phenotypes
#matrices filled with generated values
u.hat[i,]<-u
c.hat[i,]<-c
e.hat[i,]<-e
p.hat[i,]<-y
s2u.hat[i,]<-var(u)
s2c.hat[i,]<-var(c)
s2e.hat[i,]<-var(e)
s2p.hat[i,]<-var(y)
h2.hat[i,]<-var(u)/var(y)
c2.hat[i,]<-var(c)/var(y)}
Sys.time() - start #count time
```

WinBUGS code for the simulated data (chapter 2)

```
model
        {
        ## Phenotypic values
        for(i in 1 : n) {y[i] ~ dnorm(mu[i], tau2e)
mu[i] <- alpha + beta.sex[sex[i]] + beta.hatch[hatch[i]] + [id[i]]+beta.mat[mat[i]] }
        ## Location priors
        alpha ~ dnorm (0, 0.0001)
        for(j in 1 :2) { beta.sex[j] ~ dnorm (0, 0.0001) }
        beta.hatch[42]<-0
        for(j in 1 :41) { beta.hatch[j] ~ dnorm (0, 0.0001) }
        ## Maternal environmental effects
        beta.mat[75]<-0
        for(k in 1 :74) { beta.mat[k] ~ dnorm (0, tau2c) }
        ## Additive genetic values
        u[2457]<-0
        for(k in 1 :137) { u[id[k]] ~ dnorm (pu[id[k]], tau2a[id[k]])
        pu[id[k]] <- 0.5*(u[fid[k]]+u[mid[k]])
        tau2a[id[k]]<- winv[id[k]]*tau2u }
        for(k in 138 :2456)      { u[id[k]] ~ dnorm (pu[id[k]], tau2a[id[k]])
pu[id[k]] <- 0.5*(u[fid[k]]+u[mid[k]]) + corr * beta.mat[mat[k]] * (tau2c/tau2u)
        tau2a[id[k]]<- winv[id[k]]*tau2u * (1 - corr * corr) }
        ## Variance priors
        tau2e <- 1 / sigma2e
        sigma2e ~ dgamma(0.001, 0.001)
        tau2u <- 1 / sigma2u
        sigma2u <-  sigmau * sigmau
        sigmau ~ dunif (0,100)
        tau2c <- 1 / sigma2c
        sigma2c ~ dgamma(0.001, 0.001)
        corr  ~ dunif (-1,1)
        cov <- corr * (sqrt(sigma2c)) * sigmau
        ## Heritability
        sigma2p <- sigma2u + sigma2c + cov + sigma2e
        h2 <- sigma2u / sigma2p
        c2 <- sigma2c / sigma2p
        covar <- cov / sigma2p
        }
Inits list( ## Means
        alpha = 0, beta.sex = c(0,0),
        ## Variances
        sigma2e = 1, sigma2c = 1, sigmau = 1)
```

ASREML code for the sex specific GEI model Q_2 (chapter 5)

```
Genotype (by sex) by environment interaction
  tag !P
  sire !P
  dam !P
  farm !I 5
  year !I 6
  season !I 4
  age !I 4
  sexenv !I 4
  hatch !I 276
  bw
ge.ped !ALPHA !SKIP 1
ge.dat !SKIP 1 !MAXIT 20 !EXTRA 2 !CONTINUE
bw ~ at(env,1).(farm year season age season.year farm.year) at(env,2).(hatch age)
        !r sexenv.tag sexenv.ide(dam)
4 1 2
99330 !S2=0.7
103993 !S2=0.6
17990 !S2=0.5
17605 !S2=0.4
sexenv.tag 2
4 0 US !GP
0.2
0.01 0.3
0.01 0.01 0.5
0.01 0.01 0.01 0.5
tag 0 AINV
sexenv.ide(dam) 2
4 0 US
0.2
0.01 0.3
0.01 0.01 0.2
0.01 0.01 0.01 0.2
ide(dam)
```