Longitudinal aspects of the genetic analysis of reproduction traits in a modern heavy turkey line

Andreas Kranis

BSc (Athens, Greece)

MSc (Athens, Greece)

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of Edinburgh

2007

Acknowledgments

At the end of my thesis I would like to thank all those people who made this thesis possible and an enjoyable experience for me.

This thesis would not have been possible without the kind support, the valuable critiques and the probing questions of my supervisors. First of all, I wish to express my sincere gratitude to Prof. John Woolliams who guided this work and helped me to get into Mathematics and Statistics. Thanks to you John, I discover a new world and for this I ’ll be indebted. My second supervisor in Roslin, Dr. Paul Hocking honored me with his confidence in the beginning of the project and he was always assisting me whenever I was in need. Finally, the invaluable contributions from my University supervisor, Prof. William Hill, helped me to enrich this research. My supervisors’ enthusiasm and integral view on research and their mission for providing ’only high-quality work and not less’, has made a deep impression on me.

I am also indebted to Prof. Daniel Sorensen for hosting me in the Danish Institute of Agricultural Sciences in Foulum. I had the privilege to be introduced to the realm of Bayesian Statistics by him and working closely in its application in Animal Genetics. Moreover, I would like to thank Dr. Guoseng Su who helped me with the Random Regression methodology and made this a very enjoyable collaboration.

The PhD is a long journey and without the companionship of fellow students and researchers, it would be a very lonely trip. So, I am grateful to the members of Roslin and DIAS institutes for their support and their comradeship; especially to: David, Giannis, Elina, Claudia, Alex C, Alex L, Ross and Jan B. Moreover, I would like to say a big thank you to all the friends that I made during these years in Edinburgh and in Foulum for sharing with me the brighter side of this life.

I would like to acknowledge the financial support by the public benefit foundation ”Alexadros S. Onassis”. Moreover, I would like to thank British United Turkeys for providing the data, the support and the trust. Special thanks go to Dr. James Bentley and Dr. Andy Morris. Lastly, I give my thanks to Genesis-Faraday Partnership for providing all the links

with industry and help me to acquire critical transferable skills.

Finally, I would like to express my deepest gratitude for the constant support, understanding and love that I received from my parents, my aunt and my grandfather. Only with their contribution was I able to undertake this project up to the end.

There is only one person for which the last "thank you" is reserved and this is my companion in life, Chrysa. Without your love, support and affection I would be wandering hoping to meet you. I feel blessed for having find you and sharing both the nice and the bad moments with you. The world is a nicer place with you in my side :-)

Edinburgh, July 2007

Publications

- Kranis A., Su G., Sorensen D., Woolliams J. A. (2007). The Application of Random Regression Models in the Genetic Analysis of the Monthly Egg Production in Turkeys and a Comparison with Alternative Longitudinal Models. Poultry Science, **86**, pp. 470:475
- Kranis A., Hocking P.M., Hill W.G. and Woolliams J.A. (2006). Genetic parameters for a heavy female turkey line: The impact of simultaneous selection for body weight and total egg number. British Poultry Science, **47**, pp. 685: 693
- Kranis A., Su G., Sorensen D.A., Woolliams J.A.(2006). A longitudinal threshold model for genetic analysis of daily egg production in turkeys in turkeys under a Bayesian framework. Proccedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, session 7, communication 6
- Kranis A., Su G., Sorensen D.A., Woolliams J.A (2006). The application of random regression models in turkeys egg production. Proceedings of the Annual Conference of British Society of Animal Science 2006, pp. 32
- Kranis A., Hocking P.M., Hill W.G. and Woolliams J.A. (2005). Genetic parameters for a heavy female turkey line. Proceedings of the Annual Conference of British Society of Animal Science 2005, pp. 9

Abstract

The current trend in turkey breeding is to maximise meat yield as a response to market demands. Although there is evidence that the intense selection for body traits has a detrimental effect on the reproductive performance of hens, the degree of the correlated responses has not been assessed for modern turkey lines. Furthermore, little is known about the longitudinal characteristics of egg laying. The common approach for breeders is to use the total egg production over a specified period, which implicitly ignores the changes in the underlying physiology of laying. One way of providing more accurate modelling might be to account for these changes over time. Therefore, the objective of this thesis was to investigate the longitudinal aspect of the genetics of egg laying in two heavy female turkeys and explore whether considering time features of laying may improve selection efficiency. The dataset consisted of records from two commercial lines, although only one had longitudinal data. The genetic correlation between body weight and total egg number was estimated to be -0.7±0.1 and -0.5±0.1 in the two lines studied. Both estimates were highly negative and larger in magnitude than in traditional and lighter lines, suggesting that the continuous selection for growth hinders genetic progress for egg production. Heritability estimates for body weight were high (around 0.4), while being lower for total egg number (around 0.2 for egg records on which Box-Cox transformation has been applied to reduce their deviation from normality).

Since heavier birds tended to lay fewer eggs within a specific period, this implied reduced rates of lay. In order to explore this consequence of the selection for antagonistic traits, a time-to-event trait was formulated that corresponded to the days required for a hen to lay 82 eggs (the average egg production in the dataset). It was shown that the Weibull distribution could satisfactorily serve as the baseline function under a survival analysis context. So, a frailty model was constructed to perform a genetic analysis of the time trait and it was found that its heritability estimate was between that for the transformed and untransformed total egg number.

Random regression (RR) models were also considered for the longitudinal analysis of egg production. It was shown by cross-validation that these models offered lower prediction errors for missing values compared to multi-trait and repeatability models. Furthermore,

a second cross-validation strategy between RR models of different polynomial order suggested that the model with second order Legendre polynomials was the most appropriate for the specific dataset. Based on this, a liability model was developed under a Bayesian approach in order to obtain heritability profiles on a daily basis. There, it was shown in greater detail that the genetic variance changed over time: at the onset of laying it was high, but it rapidly reduced until reaching a minimum that coincided with the period of the production peak, and finally it increased towards the end of the laying period. Genetic parameters were estimated for the cumulative production of all subperiods of the laying period and used to detect efficient selection windows for which the output of selection based on partial records was better compared to the selection for the total egg production. It was indicated that periods covering the last month of the laying period appeared to be favourable on both the underlying and observed scale. In order to capitalise on the changes of genetic variance over time, a bivariate RR model for the monthly egg production and body weight was applied to detect periods where the genetic association between the traits was weaker than that between body weight and total egg number. The genetic correlation estimate tended to vary over time (it ranged from -0.4 to -0.7) but it was not possible to identify a laying stage presenting a clear advantage in combining an efficient selection for egg production and a moderate correlation with body weight.

As the genetic covariance was shown to change over time, treating egg production as a dynamic trait provided more accurate modelling of the underlying biological mechanisms. So, it may be possible to achieve higher selection efficiency when the daily laying records are allocated different weights compared to total egg production, in which equal weight is given to each day of the laying period. In conclusion, the detailed covariance structure obtained on a longitudinal basis can be used to target more effectively the selection pressure on the most informative stages of laying and thus, to maximise the output of breeding programmes.

Contents

Declaration 2

Acknowledgments 4

Publications 5

Abstract 7

1 Introduction 17

- 1.1 General introduction 18
- 1.2 Trends in turkey breeding 18
  - 1.2.1 Market demands and breeding objectives 19
  - 1.2.2 The breeder response to trends and changes in breeding schemes . 21
- 1.3 Implications of selecting for increased body weight 23
  - 1.3.1 Impact on meat quality and body composition 23
  - 1.3.2 Impact on reproductive performance 25
- 1.4 Genetics of egg production and selection for body weight 27
- 1.5 Strategies to reduce the impact of selection for body weight 32
- 1.6 Research objectives 33

2 Data Description and Handling **35**

2.1 Introduction 36

2.2 Population description 36

2.3 Dataset description 37

2.4 Data handling 38

2.4.1 Development of custom programs 41

3 Genetic Parameter Estimation **43**

3.1 Introduction 44

3.2 Materials and methods 44

3.2.1 Data description 44

3.2.2 Dealing with deviation from normality for egg production trait . . 46

3.2.3 Genetic model 47

3.3 Results 48

3.3.1 Outlier exclusion & Box-Cox transformations 48

3.3.2 Maternal effects model 49

3.3.3 Multi-trait analysis 51

3.4 Discussion 53

4 Application of Survival Analysis **59**

4.1 Introduction 60

4.2 Materials and methods 61

4.2.1 Key concepts of survival analysis 61

4.2.2 Nonparametric survival model 62

4.2.3 Parametric survival model 63

4.2.4 Frailty survival model 65
4.2.5 Estimation of genetic parameters 66
4.2.6 Defining starting point and terminating event 67
4.3 Results 67
4.3.1 Nonparametric analysis (KM curves) 67
4.3.2 Parametric model (Weibull model) 69
4.3.3 Heritability estimates 71
4.4 Discussion 72

5 Application of random regression Models 76

5.1 Introduction 77
5.2 Materials and methods 78
5.2.1 Fixed regression 78
5.2.2 Repeatability Model 79
5.2.3 Random Regression Model 80
5.2.4 Multi-trait Model 82
5.2.5 Model Comparison 82
5.3 Results 83
5.4 Discussion 87

6 Threshold Random Regression Model 91

6.1 Introduction 92
6.2 Materials and methods 93
6.2.1 Genetic model 93
6.2.2 Inferences of genetic parameters 95

6.3 Results . . . 98
6.3.1 Fixed regression . . . 98
6.3.2 Genetic parameters for the daily egg production . . . 98
6.3.3 Cumulative liability/egg production from day one to day x . . . 102
6.3.4 Cumulative liability/egg production & indirect selection . . . 104
6.4 Discussion . . . 106
6.4.1 Liability and observed scale . . . 107
6.4.2 Laying patterns . . . 108
6.4.3 Indirect selection using partial production . . . 110

7 Bivariate Random Regression Model **112**
7.1 Introduction . . . 113
7.2 Materials and methods . . . 114
7.2.1 Traits considered . . . 114
7.2.2 Genetic model . . . 115
7.2.3 Estimation of genetic parameters . . . 117
7.3 Results . . . 119
7.4 Discussion . . . 123

8 General Discussion **127**
8.1 Objectives revisited . . . 128
8.2 Key findings . . . 129
8.3 Implications . . . 132
8.3.1 Laying dynamics . . . 133
8.3.2 Capitalise upon laying dynamics . . . 136

8.3.3 Selecting for improved egg production 140

8.4 Future developments 141

8.5 General conclusions 143

List of Figures

1.1 The supply chain of poultry sector. 19

1.2 A simplified breeding scheme 22

1.3 Conceptual diagram showing the changes over time in body weight and egg number for a line selected for growth. 29

2.1 A typical example of a trapcard. Additional information is contained, including the cause of culling, the cracked eggs (marked with "X"), and the eggs laid in the floor. 40

3.1 Histograms for the untransformed (left panel) and the Box-Cox transformed (λ=3 for UK & λ=2.5 for USA) data (right panel) for total egg number in selection flocks of turkeys in the UK and USA. 50

3.2 Normal probability plots (observed probability - expected probability) for the untransformed (left panel) and the Box-Cox transformed (right panel) in the UK population. 51

4.1 Kaplan-Meir plots of empirical survival (left panel) and hazard function (right panel) stratified per year. 68

4.2 Kaplan-Meier plots of empirical survival (left panel) and hazard function (right panel) stratified per sire. 69

4.3 Comparison of survival curves expected from a Cox model and observed stratified per year against time. Lack of discrepancies between the curves suggests that the PH assumption is not violated. 70

4.4 Log-Log Kaplan-Meier survival estimates stratified for year to compare the appropriateness of datasets in which time recording started the first (left panel and 90^{th} (right panel) day of the laying period. 71

4.5 Fast laying hens tend to have higher total egg production, implying that the time-to-event trait may provide a means to assess persistency of lay (Censored records were not included in the graph. 73

5.1 Heritability profile for all models (MTM multi-trait model; RR2 Random regression model using 2^{nd} order Legendre polynomials; RR3 Same as RR2 but using 3^{rd} order polynomials). 87

6.1 Fit of the fixed regression modelling the average daily laying curve (hen-day production) over the raw data (points). 99

6.2 Skewness and kurtosis of the distribution of daily heritability on the observed scale (the curve shape was similar for both moments in the liability scale). 100

6.3 Plot of the daily heritability in the liability scale, including the standard deviation of the estimates. 101

6.4 Plot of the daily heritability in the observed scale, including the standard deviation of the estimates. 102

6.5 Surface plot of the genetic correlation between all days of laying period. . 103

6.6 Plot of the heritability estimates, including the standard deviation, of the cumulative egg production from day one up to day x ($x \leq 140$) in the liabiltiy scale. 104

6.7 Search for optimum time windows (marked as light-coloured, yellow) for which the indirect selection using the cumulative production between two days, corresponding to x and y axes of the contour plot, is predicted to be more efficient than the direct selection for the total egg number in the liability scale. 105

6.8 Comparison of the indirect selection for cumulative production beginning from day one up to day x and the direct selection for total egg number. Partial egg production starting from the first day and finishing between days 14 and 34 is the only interval when the indirect selection is predicted to be more efficient. 106

6.9 Likelihood of an efficient selection window existing when the cumulative production from day one up to day *x* ($x \leq 140$) is used for indirect selection for the total egg number in both the liability and observed scale. 107

7.1 Heritability profile for all periods. 121

8.1 Eigenfunctions of the genetic covariance function coefficient matrix. . . . 134

8.2 Eigenfunctions of the residual hen covariance function coefficient matrix. 135

List of Tables

3.1 Descriptive statistics covering all the four generations for body weight at 14 (BW14), 19 (BW19) and 24 (BW24) weeks of age for males (M) and females (F) and egg production (EGG). 45

3.2 Effects of outlier exclusion (step 1) followed by Box-Cox transformation (step 2) on the coefficient of variation (C.V.), skewness and kurtosis of egg production. 49

3.3 Log-likelihood values for all maternal effects models 52

3.4 Estimates from multivariate REML of variance components for body weight at 14 (BW14, kg^2), 19 (BW19, kg^2) and 24 (BW24, kg^2) weeks of age and for total egg number laid in a production period of 139 days (EGG) 53

3.5 Phenotypic correlations and genetic parameter estimates from the multivariate REML for body weight at 14 (BW14), 19 (BW19) & 24 (BW24) weeks of age and egg production (EGG). Heritability estimates are given on the diagonal (in bold), genetic correlations above diagonal and phenotypic below. 54

5.1 Phenotypic variance for all models and for all periods 84

5.2 Heritability for all models and for all periods 85

5.3 Estimates of the proportion of phenotypic variance explained by the permanent environment (c^2) for all models and for all periods 85

5.4 Genetic correlation coefficients between all periods 86

5.5 Phenotypic correlation coefficients between all periods 86

7.1 Genetic (co)variance over time for monthly egg production. 120

7.2 Permanent environmental (co)variance over time for monthly egg production. 120

7.3 Total phenotypic (co)variance over time for monthly egg production . . . 120

7.4 Heritability estimates for the egg production of individual periods (left column) and cumulative production up to period x (right column). 121

7.5 Phenotypic (above the diagonal - in italics) and genetic correlation (below the diagonal - regular fonts) betweem body weight at 14 weeks (BW14) and egg production of individual periods (P1 to P5). 122

7.6 Comparison of heritability estimates for three RR models 124

8.1 Heritability estimates (h^2) for the egg production in the first month of the laying period derived from a univariate animal model (h^2 with standard error (SE) for untransformed and Box-Cox transformed values) and threshold RR model (h^2 with standard deviation (SD) of the posterior distribution are given both in liability and observed scale). 138

Chapter 1

Introduction

1.1 General introduction

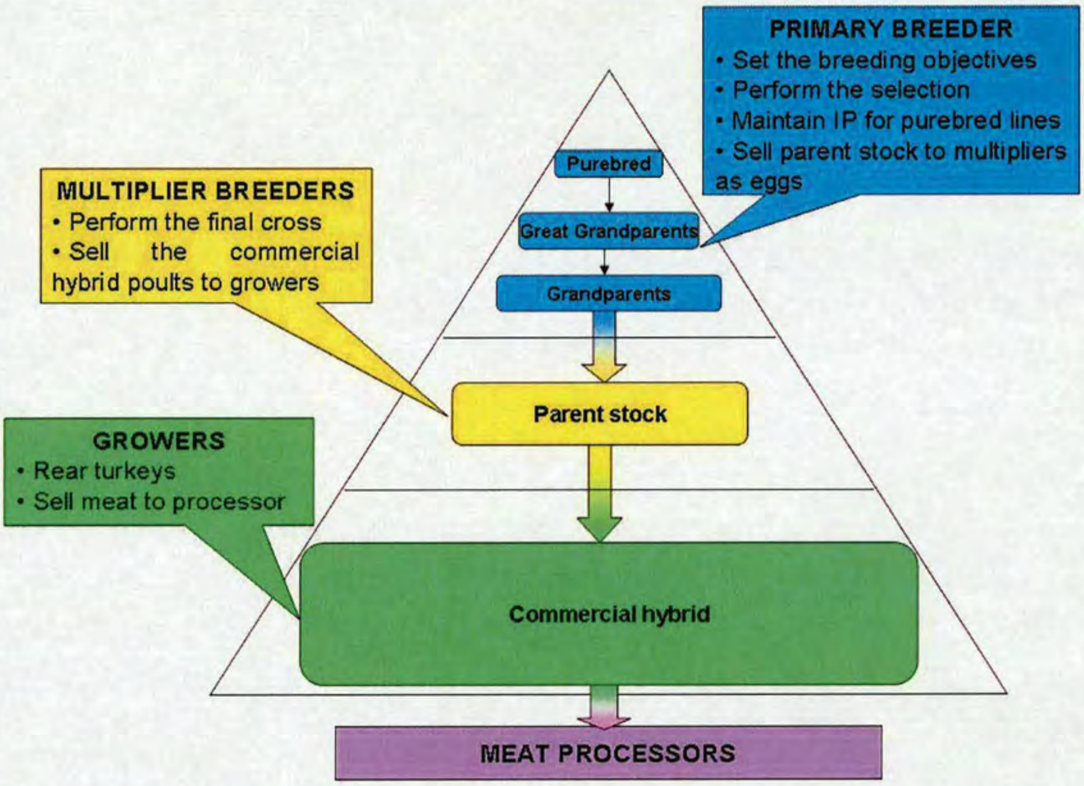
Sixty years ago, the world consumption of poultry meat was low, but since the Second World War a rapid development has been observed (Taha, 2003). The sector is characterised by a relatively constant annual growth rate (Arthur and Albers, 2003) and only recently has progress been slowed down due to the threat of the avian influenza pandemic (USDA, 2006). Although the broiler chicken industry dominates the poultry meat output, turkey and duck sectors are also significant, particularly in some countries.

The progress in the productivity of birds is unprecedented (Fairfull *et al.*, 1998). This can be attributed to many factors, such as the introduction of improved management methods and the development of novel nutrition strategies. Nevertheless, the application of modern breeding programmes accounts for 90% of the observed improvement (Reddy, 1996), highlighting the importance of well-designed breeding schemes.

This chapter provides an overview of the recent trends in the turkey industry, examines the implications of the breeders' response to market demands to focus more on growth traits and illustrates how this results in the reduction of egg numbers and disturbs laying dynamics. This was the motive to perform a pilot study on the longitudinal aspects of the egg laying in order to seek alternative strategies to improve the current selection program.

1.2 Trends in turkey breeding

The supply chain of the poultry sector consists of the primary breeders, the multipliers, the growers, the meat processors and the wholesalers/retailers (Figure 1.1). Only a limited number of breeding companies possess the breeding stock and perform the genetic improvement. The impact of the selection in purebred lines is remarkable, since one breeder hen may produce more than 1,500 tonnes of turkey meat at the commercial level. Breeding companies sell day-old poults to the multipliers, which in turn sell the final hybrids to the growers. The poultry supply chain is characterised by numerous large vertically



The size of round boxes is proportional to the population size. One breeder hen in the "purebred line" level produces 1,750,000 kg of turkey meat at the commercial level

Figure 1.1: The supply chain of poultry sector.

integrated enterprises that buy poults directly from the breeders or multipliers and sell processed turkey meat to the market. These key players have established direct communication with the breeding companies and thus the trends of the market are efficiently transferred to the breeding companies, which adapt the selection goals accordingly.

1.2.1 Market demands and breeding objectives

A significant shift in consumers' preference has been observed towards low fat and low cholesterol foods. Turkey meat, having these attributes, is consistent with modern diets and is becoming a more favourable option. Moreover, turkey meat is ideal for further processing. A large part of the total turkey meat production is destined for the food manufacturing industry. Therefore, the demand for turkey meat is increasing and the sector is rapidly developing, particularly in US (Jones, 2006).

Body weight constitutes a principal breeding objective in meat-type poultry. It is easily measured with minimum cost and is highly heritable in poultry (Emmerson, 2003). This permits a reliable selection of birds based on their individual performance. Another important growth related trait is feed efficiency, which contributes to the reduction of the production cost and decreases the environmental impact of intensive farming (Flock, 1998). However, its recording is expensive and thus, due to resource constraints, only recently it has started to be included in breeding programs in turkeys, in contrast to broiler chicken.

From an economic viewpoint, the breast muscle is the most important part of the carcass, because it represents 70% of the total revenue source and it is the most suitable as a raw material for the food processing industry. Therefore, the conformation of the bird is of paramount importance to breeders in order to maximise the yield of the breast muscle. The selection criteria used to assess the bird capability for breast meat yield include phenotypic assessment of the keel and shank length and the width of the breast muscle (Buss, 1990) expressed as subjective conformation scores by experienced personnel. All these traits are highly heritable (Buss, 1990; LeBihan-Duval *et al.*, 1998) and positively correlated with body weight (Wilkiewicz-Wawro *et al.*, 2003). This classification system has been proven effective, primarily for fleshing (Emmerson, 2003). Nevertheless, the lack of objective measurements increases the variability of commercial lines originating from different breeding companies (Nestor *et al.*, 2001). A more direct approach is to incorporate information from the carcass composition from half-sibs in the selection index. Additionally, measurements on live animals can be taken with ultrasonic devices. The extra cost from the application of such methods can be justified in broiler breeding, but for turkey breeders the cost-benefit ratio may not yet be persuasive, because the turnover of this industry is significantly smaller. Nevertheless, obtaining an insight to the body composition of breeding hens may provide significant information and this is why recently an investment on introducing this technology is under way in turkeys.

The breeding objectives include also reproduction traits. Both breeding companies and multipliers are interested in the maintenance of a satisfactory egg production in order to replace the lines and sustain the efficiency of the production system. Therefore, the traits that turkeys are selected for are fertility and hatchablity, as well as the total egg number within a defined laying period.

1.2.2 The breeder response to trends and changes in breeding schemes

The market demand for heavier birds dictates the need for the breeding companies to meet this objective (Douglas, n.d.). This goal appears to be feasible, since the genetic variance for growth traits remains high, suggesting that the selection limits have not been reached yet for them (Havenstein, 2003). However, changes in the structure of the breeding programmes may be required in order to achieve a rapid increase in the body weight of the commercial hybrid.

The traditional structure of the breeding programmes involves a number of purebred lines that are crossed in order to combine the desirable traits by maximising the benefits from the heterosis. An overview of the applied system is given in Figure 1.2.

The purebred lines remain the property of the breeding company and constitute the nucleus stock in which the genetic selection is performed. They are distinguished in two categories, male and female lines. The former, also called sire lines, (A & B in Figure 1.2), are selected for body weight and body conformation, especially for wide breast muscle, and their progeny gives the male for the grandparent stock. The latter, also called dam lines, (C & D in Figure 1.2), were tradionally selected for reproductive traits, such as the total egg production, and they give the female for the grandparent stock. This difference in selection criteria is reflected to their body weight; male lines are heavier and larger, and female lines are lighter and smaller in size. The progeny of the purebred lines is the great-grandparent stock, which is also selected for the target traits and it is used mainly to multiply the line to obtain the population size required to produce the next stock. This is the grandparent stock, which is sold by the breeding company to the mul-

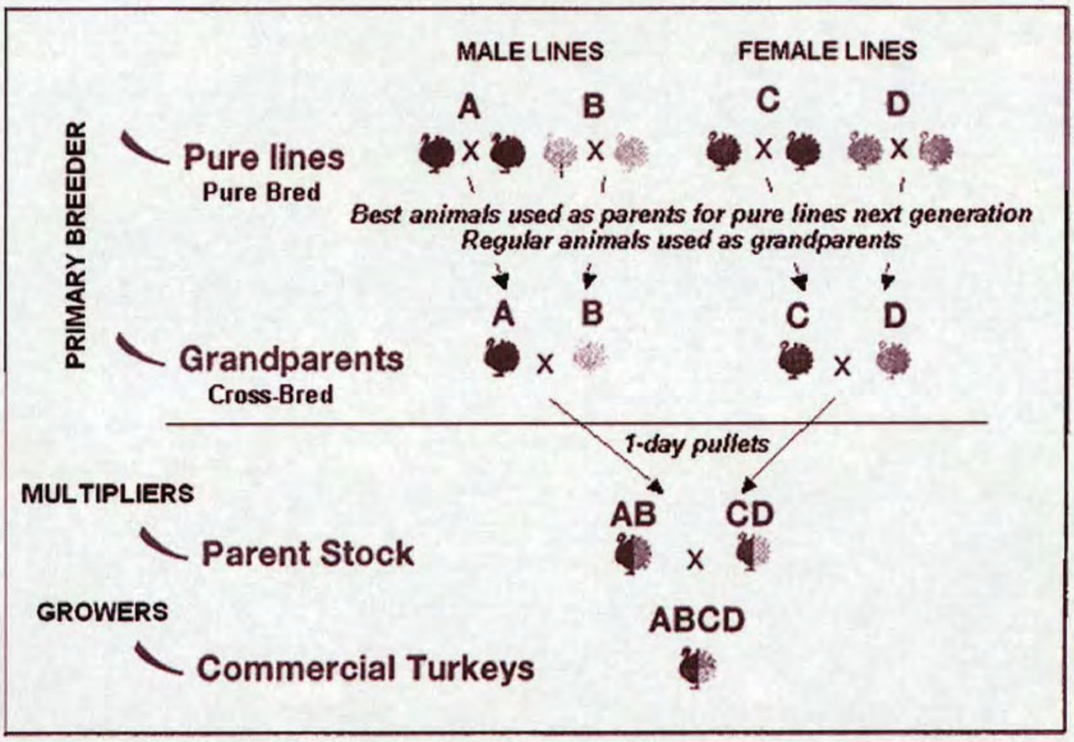


Figure 1.2: A simplified breeding scheme. The crosses above the line are performed by the primary breeding company and the crosses below the line by commercial multipliers (modified from www.hybridturkeys.com).

tipliers or to the integrated production companies as hatchable eggs or one-day poults. These birds, being hybrids, have superior performance for both growth and reproduction traits, but they cannot be used to regenerate the original lines, effectively protecting the ownership of the elite genetic material of breeding companies (van Arendonk and Bijma, 2003). Grandparent stock is crossed again to obtain the parent stock (AB and CD). The final multiplication gives the commercial hybrids (ABCD) that are distributed to farms for growing, slaughtering, processing and consumption. In total, four generations separate the purebred lines and the final hybrids.

This multiplication scheme offers to the breeding companies the flexibility to adapt quickly to the market demands by choosing which lines participate in the final crosses. For this reason they maintain several purebred lines that are selected for different breeding objectives. Although the reproductive traits from the female lines are important to multipliers, it is the growth traits, inherited from the male lines, that are crucial to growers. Due to competition, the challenge for breeders is to achieve a rapid increase in the body weight of turkeys. So, it can be inferred that the best strategy to achieve this goal is to put selection pressure for growth not only to the sire but also to the dam line. This change in the selection process only recently has been applied. However, this could have several implications, which are discussed in the next section.

1.3 Implications of selecting for increased body weight

When extra selection pressure is put on a trait in order to improve productivity, unwanted correlated responses in other production traits may occur. The intense directional selection for body weight and conformation causes modifications to the body composition and to the hormonal background of turkeys. The impact of this change is multiple: modern broilers and turkeys are known to suffer from skeletal defects, particularly in legs (Whitehead *et al.*, 2003); reduced immunological responses (Bayyari *et al.*, 1997) that could provide an explanation why heavy lines are more susceptible to diseases such as fowl cholera (Li *et al.*, 2001); and decreased adaptability to environmental conditions (Rauw *et al.*, 1998). These problems may contribute to the reduced viability that is observed in heavy turkeys (Havenstein *et al.*, 2003). The following sections focus on the impact of the selection for increased weight on body conformation and how this is related to the meat quality and reproductive performance.

1.3.1 Impact on meat quality and body composition

Meat quality is a complex trait involving physiological aspects of muscle and fat tissue; organoleptic properties, such as the colour and tenderness; and factors that are related to the processing capability, such as the water-holding and emulsifying capacity. This section focuses on the impact of selection for growth on muscle and adipose tissues and how this relates to the meat quality. It should be mentioned that the majority of the published results refer to broiler chicken, so caution is suggested when making inferences for turkeys. A comparison of genetic parameter estimates for the same traits in broiler chicken (LeBihan-Duval *et al.*, 2001) and turkeys (Remignon and LeBihan-Duval, 2003) revealed quantitative differences between the two species.

There is evidence that selection for body weight affects the morphology of the muscle fibres in turkeys (Swatland, 1989) and increases the levels of plasma creatine kinase, which is associated with muscle damage (Wilson *et al.*, 1990; Hocking *et al.*, 1998). These muscle abnormalities are more frequent in fast-growing turkeys, as well as the increased occurrence of pale, soft and exudative turkey meat (PSE). The latter constitutes a major problem for the turkey meat industry, since it may affect as much as 40% of the total production destined for processing (Owens *et al.*, 2000). Whereas the causative mutation of PSE meat in pigs has been detected (Otsu *et al.*, 1991), in turkeys the mechanism seems not to be the same and contradictory results exist. Fernandez *et al.* (2001) reported that no significant differences in post-mortem metabolism were found between fast and slow growing turkey lines. However, reports from long-term selection experiments suggested that a genetic association of PSE occurrence and growth rate is likely (Velleman and Nestor, 2004). In conclusion, it appears that selection for rapid growth may increase the output of the total meat production, but it has a negative impact on the quality of the turkey meat.

The adipose tissue contributes to the flavour and tenderness of the meat. However, fat in excess of requirements decreases the feed efficiency, increases the cost of rearing and reduces the processing capacity of the meat, because fatty carcasses are more perishable. For these reasons, an indirect selection is applied using the body conformation and feed conversion traits to reduce the fat. A direct approach would include ultrasonic measurements on living birds or information from slaughtered half-sibs. Excess fat is stored in the abdominal cavity and constitutes an indicator of the overall fat level. Indirect selection for growth seems to have conflicting effects on the fat content. On the one hand, the genetic correlation between body weight and abdominal fat was estimated to range between 0.2

and 0.5 for broiler chicken (Chambers, 1990; LeBihan-Duval *et al.*, 1998, 1999) and to be 0.25 in turkeys (Chapuis *et al.*, 1996). On the other hand, breast muscle yield was found to be negatively correlated with abdominal fat in broilers, but the association appeared weak, being equal to -0.1 (LeBihan-Duval *et al.*, 1998, 1999; Berri *et al.*, 2001). However, another possible factor that may also be involved is the subjective scoring for optimum body conformation, given by the breeder. The desirable conformation is to have birds not only with wide breast but also with a developed and round rear part, so a selector having these criteria may favour birds with increased abdominal fat. In conclusion, it appears that selection for body weight increases the fat content in heavy turkey lines (Buss, 1990).

1.3.2 Impact on reproductive performance

The reproductive performance represents a combination of a plethora of traits that are affected by both genetical and environmental factors, such as the age of sexual maturation, sperm properties, fertility and hatchability. The selection for body weight and conformation causes problems in the reproductive effectiveness of both male and female birds. An obvious consequence is that male turkeys are so heavy that natural mating is impossible and thus artificial insemination is essential (Rauw *et al.*, 1998). Furthermore, negative genetic correlations are observed between semen yield and fertility and body weight (Nestor, 1976, 1977b). The effects appear analogous to chicken, so this section provides an overview of the implications in relation to the body conformation and composition.

The selection for wide breast muscle is associated with increased egg production in ducks (Farhat *et al.*, 1998). However, selection for low body fat may have the opposite effect. Thus, the extreme leaness may result in a delay of sexual maturity, since a critical fat content is associated with the initiation of egg production in broilers (Katanbaf *et al.*, 1989). It should be noted though that the opposite extreme body condition, excessive fatness, was also linked to the incidence of reproductive problems, as described below (Fairfull *et al.*, 1998).

The intense selection for body weight is associated with increased incidence of oviduct prolapse (Hocking, 1993). This is a severe failure of the reproductive system that leads to culling and can result in losses up to the 10% of the flock (Buchanan *et al.*, 2000a). The aetiology remains unknown, although it was postulated that in heavy lines ovulation is initiated before the oviduct reaches a mature stage and therefore these hens are susceptible to prolapse (Melnychuk *et al.*, 1997). Nevertheless, in later studies it has been suggested that prolapse is likely to be physiologically associated with the effect of oestrogen on collagen breakdown and an interference of growth factors has been speculated (Buchanan *et al.*, 2000a). The latter may suggest a plausible pathway to explain the correlation between the selection for increased body weight and prolapse incidences. Tablante *et al.* (2000) associated the obesity with the prolapse and highlighted the negative impact of the latter on welfare, since it invokes the vent-pecking form of cannibalism.

Increased body weight is considered to be a predisposing factor for multiple ovarian hierarchies (Melnychuk *et al.*, 1997; Hocking and Bernard, 1998). Although in a normal ovulation cycle only one follicle develops, it is possible that additional follicles grow simultaneously. This results in an internal ovulation or double-yolked eggs that cannot be hatched (Buchanan *et al.*, 2000b). The impact of photostimulation is believed to mediate the incidence of multiple hierarchies by postponing the initiation of laying (Hocking *et al.*, 1988, 1992). This provides additional evidence that rapid growth is associated with poor reproductive performance. Furthermore, multiple hierarchies were also linked with excessive fat due to long-term selection for body weight at an early age (Hort *et al.*, 2000), suggesting another underlying mechanism. Growth factors appear to be involved in the regulation of follicular development (Decuypere *et al.*, 2002). These factors may interfere with regulation of the secretion of reproductive hormones. It was suggested that low levels of oestrogen affect the ovarian function and thus are associated with the multiple hierarchies (Buchanan *et al.*, 2002). An additional implication of the increased body weight seems to be broodiness, to which heavy turkey lines are also prone (Elhalawani and Rosenboim, 1993).

In conclusion, the increased body weight is associated with a series of linked malfunctions of the reproductive system, which contribute to the poor laying performance of heavy turkey lines. Although the exact mechanisms have not been fully understood, the evidence suggests that the selection for rapid growth may lead to multiple ovulation and premature initiation of laying, which in turn results in decreased reproductive performance. Modern turkey lines have the genetic potential to rapidly gain body weight if fed *ad libitum*, albeit this would be detrimental for their reproductive performance (Kerr *et al.*, 2001). It can be postulated that the impact of selection for growth on body composition and particularly on fat deposits may be significant, perhaps mediated through growth factors or other components of the endocrinological system. Nevertheless, further studies are required to investigate this hypothesis.

1.4 Genetics of egg production and selection for body weight

One of the most used traits to assess the reproductive performance of hens is the total number of eggs laid within a specific period. The reason for its popularity is that total egg production is the most obvious trait, offers a summary value and is correlated with other important egg-related traits. According to the selection on egg or growth traits, poultry industry has specialised in two branches: a laying and a meat sector. The comparison between lines in which the emphasis is given either to egg or meat yield is interesting but requires caution due to the different genetic backgrounds involved. The same applies also for comparisons between species. Nevertheless, the number of published results for the genetics of egg production in turkeys is limited. The vast majority of studies have been performed on table-egg laying chicken, which hampers the comparison of results.

A large variability in the heritability estimates of total egg number is observed in turkeys (Marks, 1990). Estimates range from 0.02 (Whitson *et al.*, 1944) to 0.61 (McCartney *et al.*, 1968). Using these estimates to find a consensus value is not applicable for many reasons. First, the majority of the published results refer to analyses performed up to forty years ago, using traditional or experimental turkey lines whose genetic background has very low similarity to the modern ones. Second, these studies employed classic estimation methodology, such as the ANOVA that is not robust with large and unbalanced datasets and to effects of selection, and leads to biased estimates of variance components (Hofer, 1998; Gianola, 2000). For an historical overview of the estimation methods for genetic parameters in poultry see Besbes and Ducrocq (2003). These results are indispensable for tracing back the evolution of selection in turkey, but have limited value for the evaluation of modern turkey lines, especially for those that have been intensively selected for growth. With the exception of those of Chapuis *et al.* (1996), there are no estimates for the genetic parameters of egg production in contemporary turkey lines. In conclusion, the absence of up-to-date estimates dictates the need to obtain them for modern lines employing current methodology, such as the restricted maximum likelihood framework (REML) that has become popular for obtaining point estimates (Szwaczkowski, 2003).

The situation is analogous for the estimates of the genetic correlation between egg production and body weight. Estimates rely heavily on the genetic background and selection history of the populations examined. Hence, a compilation of older results given by Arthur and Abplanalp (1975) on light-type birds suggests a weak correlation of -0.1. However, it can be postulated that the genetic gain accumulated after consecutive generations of selection for body weight strengthens the genetic correlation. Such evidence arises from long term selection experiments for either body weight or egg production extending over more than thirty generations. Nestor *et al.* (2000) reviewed the progress made and they concluded that the magnitude of correlation changed along with selection for body weight. Since the start of the experiment body weight at sixteen weeks of age doubled and egg production reduced significantly, particularly at the first and last ten generations. In contrast, in a line selected on a long-term basis for egg production, egg production had almost doubled within the experiment period, while body weight had significantly reduced (Emmerson *et al.*, 2002). Both experiments suggested that the progress for the selected character is constant, while the correlated response for the other one changes over time. Results reported not only by Nestor *et al.* (2000) and Emmerson *et al.* (2002), but also by LeBihan-Duval *et al.* (1998) lead me to suggest a general form of responses shown in Figure 1.3.

These observations suggest that intense selection for body weight has a detrimental effect on egg production. However, the magnitude of the effect for modern heavy-type turkey

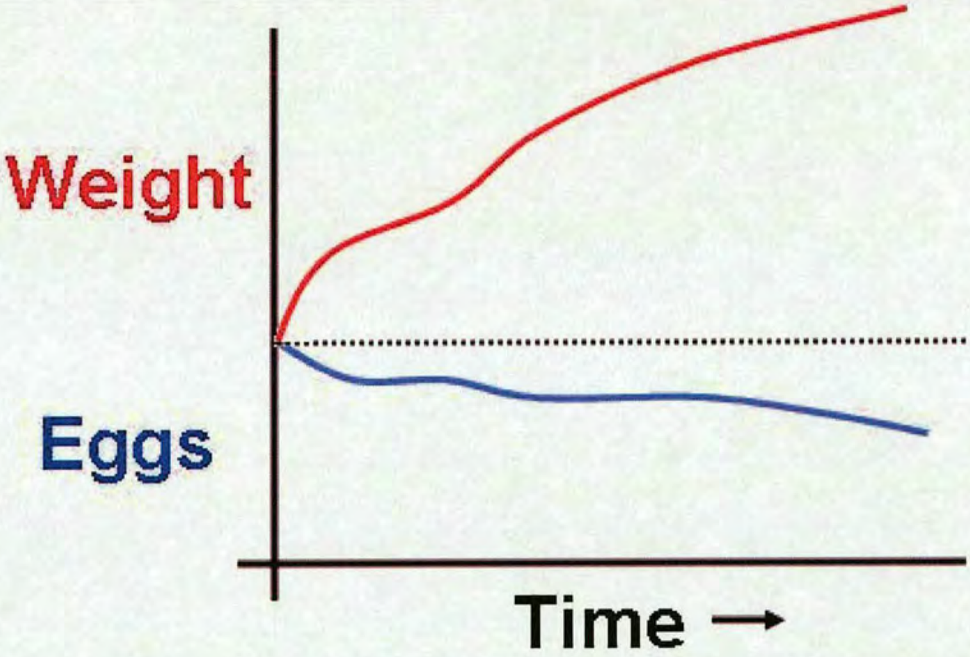


Figure 1.3: Conceptual diagram showing the changes over time in body weight and egg number for a line selected for growth.

lines that are simultaneously selected for body weight and egg production is not known as no data are available. Nestor (1977a) reported that progress in body weight may be feasible without a decreased egg production when a selection index is used. However, this study referred to a light line that was selected only for nine generations and so, the correlation between the two traits was likely to be low, as discussed in the previous paragraph. Breeders may use a selection index to improve both traits, but it is likely that the improvement would decline over time, particularly if the genetic correlation is strong. In a tandem selection scheme Nestor (1985) concluded that it was not feasible to maintain body weight at the same level and increase egg production after selecting for body weight for eleven generations. In conclusion, it can be postulated that the relationship between body weight and egg production is dynamic depending heavily on the selection history, but in order to evaluate this hypothesis additional results are required from studies performed on modern turkey lines.

Another reason for studying the association between growth and egg production is the effects of continuous selection for body weight on the reproductive performance of turkeys. As described above, the incidences of multiple hierarchies are increased in heavy turkey lines. At the phenotypic level this is translated into a reduced number of consecutive days in which a hen lays eggs, resulting in smaller clutches. Thus, irregular ovposition results in less eggs laid within the production period. The other negative consequence is that the number of hatchable eggs is reduced due to the increased incidence of double-yolked eggs and eggs with shell defects. In broiler chicken it was found that defective eggs may account for the 17% of the total production (van Middelkoop and Siegel, 1976; Anthony *et al.*, 1989). This phenomenon is also known as EODES, an acronym for erratic oviposition and deffective egg syndrome.

An additional implication of increased incidence of erratic laying patterns in heavy turkey lines is that the prediction of their production capability is hindered. Various mathematical models have been suggested to describe the laying curve of hens (for a review of older models see Fairfull and Gowe (1990)). The main characteristic of egg laying is that it is a multiphase process comprising three stages: an initiation period, in which the hen-day production is rapidly increases until it reaches a peak; a stabilisation period, in which the daily production remains constant; and finally, a decreasing period, in which the rate of laying drops. It appears that older models failed to account for all these features, since they did not fit data satisfactorily (Fairfull and Gowe, 1990; Anang *et al.*, 2001). A recently proposed model by Grossman and Koops (2001) appears to offer a better fit. It consists of two components, the one accounting for the increasing phase of laying and the other for the decreasing one. This model was intended to describe individual laying curves and consequently it fails to model average performance. Moreover, it depends on many individual-specific variables that increase complexity, although it accounts for the persistency of laying extending the concept introduced by Grossman *et al.* (2000). Nevertheless, all these models have been devised to describe the production curve in table-egg chickens, which have been intensively selected for optimised and consistent laying patterns in order to maximise the overall yield. If most of them do not offer a robust performance on regular laying patterns, it is unlikely that they would be suitable to describe the laying curve of lines from another species that has been selected for growth.

The dynamic character of laying is not only relevant to modelling of the egg production, but it is also important for the more detailed genetic evaluation of hens. The use of the total egg number as a selection criterion conceals the patterns of laying and does not enable information to be obtained on the changes in genetic variance during the production period. An approach to address this problem is to investigate the cumulative production over parts of the whole laying period and examine the covariance structure between them. This can be achieved by using a multi-trait (Anang *et al.*, 2000) or a repeatabilty model (Szwaczkowski, 2003). An alternative approach is to exploit the potential of covariance functions or the equivalent random regression models to accurately model the changes of covariance over time (Kirkpatrick and Heckman, 1989; Kirkpatrick *et al.*, 1990; Meyer and Hill, 1997; Meyer, 1998). The experience from dairy cattle (Werf *et al.*, 1998) suggests that the random regression models could offer an alternate method to analyse egg production under a longitudinal prospective. Nevertheless, poultry breeding has yet to benefit from their application. Only one published study made use of them in reporting results for a light, table-egg laying chicken line (Anang *et al.*, 2002). As it has been illustrated in the preceding section, the impact of intense selection for body weight has a highly significant effect on the overall laying profile. So, this warrants research on modern heavy lines in order to investigate further the longitudinal aspects of the egg production.

Nevertheless, it should be noted that the erratic laying patterns of laying, due to variable clutch size across hens, add a layer of complexity in the effort of modelling egg production over time. Moreover, daily egg production is an "all-or-none" character that introduces further complications in the statistical analysis of the data. Across all the thesis these issues are discussed in more details and especially in chapters 3 and 6, where different approaches are investigated to overcome the aforementioned points.

1.5 Strategies to reduce the impact of selection for body weight

Modern turkey lines have the genetic potential to rapidly gain body weight. Although this is desirable for fulfilling the commercial objective, it hastens the onset of lay, results in reduced overall yield and causes a plethora of reproductive problems, such as prolapse and multiple hierarchies. Two potential strategies can be suggested to reduce the negative consequences of the increased body weight. The first one focuses on the regulation of the growth rate and has been successfully applied in broiler chickens, while the second one involves the use of sophisticated genetic tools to obtain an insight of the egg production over time.

An effective husbandry method to decelerate the growth rate of the breeding stock is to impose feed restriction. The positive effect from the application of this management strategy has been established in broilers and it has been shown that egg production of the parent stock of broilers is dramatically improved in comparison to *ad libitum* fed birds (Klein-Hessling, 2003). However, the benefits from the application of feed restriction in turkey breeder hens appear to be limited. Although the prevalence of multiple ovulations is decreased, the overall laying performance is not improved (Hocking, 2003). A side-effect of the feed restriction is that it affects the welfare of birds (Renema and Robinson, 2004) and thus it might not be compatible with future rearing regulations. Other management methods may involve the application of alternative lighting schemes, because the initiation of laying in turkeys is regulated by the photoperiod, but further research is required.

The second approach is to examine alternative methods to select for reproductive performance. As it was illustrated in the previous sections, selection for growth triggers a concatenation of responses that have a detrimental effect on egg production and affect the dynamics of laying. Nevertheless, information on the longitudinal pattern of egg production is not adequate. So, the first step would be to exploit the potential of the statistical tools available in order to obtain an insight into the genetics of egg production over the laying period. This would allow focus on the patterns of laying and the seeking of alternative traits that would be more relevant to the underlying mechanisms that govern the association of growth and reproduction in turkeys. The final goal would be to optimise the combined selection for growth and reproduction.

1.6 Research objectives

The current trend in turkey breeding is to place great emphasis on body weight and conformation in order to respond to market demands. However, intense selection for growth has multiple negative effects on other important traits, including meat quality, and particularly to the reproduction performance. Moreover, the reduced reproductive performance may result in reduced size of families and this could increase inbreeding (by decreasing the effective population size) and reduce the number of available progeny for performing half-sib test (e.g. for meat-quality traits). This implies that the sustainability of such a selection practise is in question and thus, it stresses the need to obtain more detail on the genetics of egg production. So far, the dynamic character of egg laying was neglected since egg production was treated as a summary measurement for the whole period, but statistical developments have provided a framework to perform a detailed longitudinal study. However, there is a lack of published results and especially for modern turkey lines. Therefore, an exploratory research is required to shed light on the dynamics of laying and focusing on the accurate modelling of the genetic covariance over time.

In order to carry out this research egg records collected on a longitudinal basis from a modern turkey line that was selected both for growth and reproduction traits were used. A detailed description of the dataset is given in Chapter 2. The research objectives were defined as follows:

- To obtain genetic parameters for the population used in this study and to investigate the strength of the genetic association between growth and reproduction traits in a modern turkey line selected for body weight, conformation and egg production

(Chapter 3).

- To examine the time related patterns of egg laying by using survival analysis (Chapter 4).
- To investigate the application of random regression models in the genetic analysis of egg production of turkeys; to compare results from a multi-trait (MTM) and a repeatability (REP) model using the same dataset; and in order to assess the efficiency of each model for predicting missing values, used a reduced dataset to predict values of the deleted records for each alternative model (Chapter 5) .
- To estimate genetic parameters for the daily egg production over the whole laying period in turkeys using a longitudinal threshold model under a Bayesian framework; and second, to explore the possibility of using the information obtained in order to investigate alternative selection strategies for the genetic improvement of laying turkeys (Chapter 6).
- To estimate the strength of the genetic association of the monthly egg production with body weight by performing a bivariate analysis employing RRM (Chapter 7).

Overall, the aim of this study was the investigation of the longitudinal character of egg laying and whether a genetical analysis taking account of the time-related properties of egg production can enhance the output of breeding programs for improving egg production in turkeys.

Chapter 2

Data Description and Handling

2.1 Introduction

The population used in this study consisted of a modern female line that was simultaneously selected for growth and reproduction traits. This line was purebred, owned and developed by a breeding company, and it represented the breeders' response to the market demands for increased body weight. So, birds of this line were heavier than those of the regular female lines and at the same time, they appeared to maintain a better reproductive performance than the male lines. However, it seems that negative effects on egg production have started to emerge and thus, the sustainability of this breeding scheme is jeopardised.

A detailed description of the population studied and of the performed undergoing selection follows. Records for body weight and daily egg production were available for five generations, but the dataset demanded special treatment in order to be analysed, so a description of the data handling procedures is also included. Brief details are also given for the software developed specifically for this project.

2.2 Population description

Each year poults were born in hatches and were transferred to rearing farms under an all-in/all-out policy to maximise biosecurity. So, birds belonging to the same hatch formed a contemporary group. Individual growth was measured and selection performed at two ages. The first selection was performed at the age of fourteen weeks. Mass selection was applied using body weight and conformation as criteria and additionally a walking test was performed to assure that the heavy birds were free from leg deficiencies. Poor performers were culled. The selection intensity was higher for males than females (for details see Table 3.1). The second selection stage was performed at twenty-four weeks of age. The criteria used were the same as in the previous stage and birds, according to their performance, were either culled or selected to be used as parents for the next generation. The number of females selected as layers was fixed for each year: in total 480 hens were used as layers

At 29 weeks of age, birds were transferred to the laying farms. Hens were assigned to pens. Males were kept in separate pens, because in turkeys the mating is performed only by artificial insemination. Hens were photostimulated for a period of two weeks in order to initiate egg laying. Hens laid in individual trap-nests and their production was recorded on a daily basis for the whole laying period, which extended over twenty weeks. A preselection was performed after six weeks of egg production by identifying top performing hens based on their so far performance and the final step was at the tenth week of the laying period. A selection index was employed, including both individual performance on body weight and egg production and pedigree information for these traits. Since this was a female line, additional reproductive traits, such fertility and hatchability were also considered. Top performing hens were selected as parents for the next generation of the purebred line, while hens that scored less satisfactorily were assigned as parents for the great-grand parent stock, as illustrated in Figure 1.2.

The line was fully pedigreed. The family structure was hierarchical, with dams nested within sires and thus only maternal full-sibs and paternal half-sibs existed. On average each sire was mated with three dams and the average family size per dam was approximately 30 offspring. Matings were specifically designed in order to prevent inbreeding. The pedigree available for this study extended over five consecutive generations for the UK population (see next section for detailed description). The average rate of inbreeding (ΔF) in this populaton was less than 1% and only a handful of birds belonging in the last generation had an inbreeding coefficient equal to 25%. These estimates were obtained using the package *matvec* (Wang *et al.*, n.d.).

2.3 Dataset description

The dataset consisted of two sets of traits, body weight at various ages and daily egg production over the whole laying period. For growth-related traits, records for the body weight at the age of fourteen, nineteen and twenty-four weeks were available for four generations. In total weight data for two populations were available: the first one was situated in UK and the second one in the USA. Further details on the analysis of these traits are given in Chapter 3.

Egg production, which was recorded on a daily basis for twenty weeks for five generations. Summary data for the production over 140 days were available for both the UK and USA populations and these were used for the analysis in Chpater 3. However, daily egg data were available only for the UK population. In total the production records of 2400 hens were available for the longitudinal analysis. The dataset consisted of a sequence of 140 binary records (0: no egg; 1: egg laid) each one corresponding to a specific day of the whole production period. Eggs were collected daily because if they were left over, broodiness can be initiated, which results in the interruption of egg laying. Some hens may also cease laying for a prolonged period, so in this case they were put in the nest in order to ensure that eggs were not laid in floor. If a hen died during the laying period, the rest of its records after the date of the culling were treated as missing values. The mortality was found to be, around 5% of the initial population (hen housed). Cracked eggs were treated as no eggs, because they could not be hatched. Eggs laid on the floor were also excluded, since they could not be assigned to a hen. In conclusion, the trait considered in this analysis was the trap nested production of hatchable eggs. The analysis of these data is the topic of Chapters 3, 4, 5, 6 and 7.

2.4 Data handling

Body weight data were contained in an Excel spreadsheet format, so it was simple to import them into the database. Each bird had a unique identifier. Nevertheless, this was a large string of numbers and letters, so it was converted into a simpler format, as described in the next section.

The daily egg production was recorded on paper sheets, called trapcards. A typical example is presented in Figure 2.1. Each trapcard contained information on the pen and the farm. Hens were identified using three-numbered tags, which were reused each year, so it was necessary to link these tags with the unique identifiers in the pedigree file and the body weight data. Records for daily production were noted in a calendar-styled fashion. Incidences of an egg were denoted as one, while a blank marked the "no egg" status. An "X" denoted cracked eggs, which were not considered in the analysis. If a hen was culled, the date and the cause was marked. For instance, in Figure 2.1, the hen tagged as 626 was removed on the 15^{th} of January due to prolapse. Culling decisions were also made in case of leg problems and for non-identified causes. Each day, the number of eggs laid on the floor was also noted, although these were not included in the analysis.

Trapcards were handwritten and thus it was necessary to digitise them before further processing. Due to poor legibility, the records were typed into a computer file. This was undertaken in Rothamsted Institute under the supervision of Professor Robin Thompson. The text files produced were imported to the database program in order to create a master database that contained all the available data for each bird included in this analysis.

This was the starting point to generate a number of egg-related traits that were used in the subsequent analyses. Hence, the first trait was the total number of egg laid within the 140 days available. This was generated by simple summing of all the records for each hen. The total egg number was used in the estimation of the genetic parameters as described in Chapter 3. A second trait was the monthly egg production, which corresponded to the sum of the eggs laid within four week periods. This trait was used to investigate the application of random regression models in analysis of the genetics of egg laying, described in Chapter 5. Another trait was the time that each hen needed to lay 82 eggs, which was the grand mean over the five generations. The rationale of this analysis is given in Chapter 4. There it is also discussed how additional traits can be generated using this database on the laying patterns that can be exploited to investigate the dynamics of the egg production.

In order to perform the analyses described in the next chapters a number of tools was used. Initial statistical analyses were perfromed using GenStat (Genstat, 2006), SPSS (SPSS,

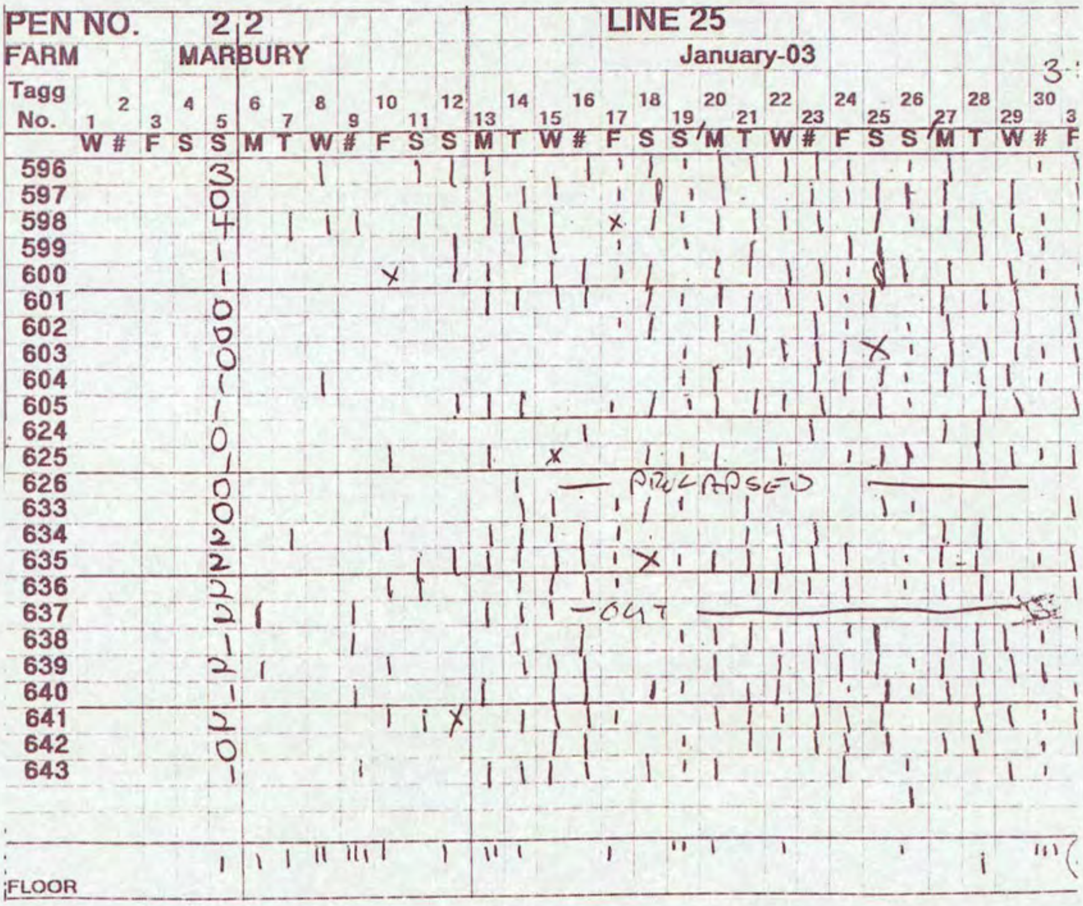


Figure 2.1: A typical example of a trapcard. Additional information is contained, including the cause of culling, the cracked eggs (marked with "X"), and the eggs laid in the floor.

2006), Stata (StataCorp, 2006) and R (R, 2006). Pedigree structure (as presented in Figure **??**) was visualised with the Pedigree Viewer program (Kinghorn and Kinghorn, 2006). The estimation of variance components estimates was performed using ASReml (Gilmour *et al.*, 2001) and DMU (Jensen and Madsen, 2006). Project specifications included also the necessary tools to develop custom programs under the Windows and Unix platforms (in both Unix flavours Linux Fedora Core 4 and Mac OS X 10.4) and visualisation tools such as the gnuplot. A description follows below.

2.4.1 Development of custom programs

The implementation of the programs was performed using Fortran 95 and Perl v5.8.6. Furthermore, SQL queries were used to extract data from the database and Unix scripts were written to perform a series of data handling related tasks.

The coding scheme used by the company for the pedigree employed large tags for each hen, so the first task was to write a program to simplify the coding in order to fix potential compatibility issues with the software used in later stages. This simplified tagging scheme was used to link the information on the summary and daily production records. In conclusion, a new coding system was applied that allowed the unique identification of birds and a link to their production data. The typing of the data on the trapcards was a painstaking task and minor errors were unavoidable, so the second task was to check for inconsistencies between the handwritten and the digitised trapcards. Having created the database, several scripts were written to generate the necessary datasets for the analyses that are described in the following chapters.

The output of an analysis employing random regression (RR) models is the covariance matrix of the regression coefficients. Further manipulation is needed to obtain the profile of the genetic parameters over time, so for this task specific programs were also developed. A simplistic solution would be to perform the computations with a mathematics package, but a suite of tools was used instead to accommodate a series of features. Hence, this program provided a simple interface to compute and visualise the profiles of the genetic parameters. Additionally, it was possible to estimate standard errors (SE) for the estimates of genetic parameters obtained using a REML framework. An outline of the estimation algorithm can be found in Chapter 5. The estimation of SE was also extended to be used in a bivariate RR model (see Chapter 7). Because Perl does not provide built-in functions for matrix operations, a custom library was written for this purpose.

Additionally, programs in Fortan were written to process DMU output for the anlysis described in Chapter 6. There, the probabilistic framework provided by the Markov Chain Monte Carlo (MCMC) method allowed estimation of statistics of the posterior distributions of the genetic parameters and the generation of novel traits, such as the cumulative production between any two time points of the laying period. So, the programs developed summarised the results from the MCMC realisations in order to generate the key figures as these can be found in Chapter 6.

Chapter 3

Genetic parameters for a heavy female turkey line: The impact of simultaneous selection for body weight and total egg number

3.1 Introduction

As it has been discussed in Chapter 1, the selection for growth causes multiple negative implications on reproduction traits. However, in case of a simultaneous selection for both body weight and conformation and total egg number, the magnitude and intensity of the correlated responses on either traits has not been investigated thoroughly for modern turkey lines. Inferring the estimates from the literature appears not to be an option, since only a handful of published estimates exist and none of these studies has been performed on a line in which combined selection is applied.

Figure 1.3 presented a conceptual diagram for the evolution of body weight and egg production when growth traits have been selected over a prolonged number of consecutive generations. This graph represented a compilation based on evidence from published longitudinal studies. Nevertheless, the validity of this hypothesis can be tested using data from the purebred lines of a major breeding company.

Therefore, the objective of this study was to investigate the effect of dual-purpose selection in two experimental populations by estimating the genetic parameters for each. Both populations derived from the same heavy female line that was simultaneously selected for body weight and egg number, but they were located in different countries, the first in the UK and the second in the USA. Various mixed models applying multivariate REML were used in the analysis to investigate maternal effects, and transformations were employed to reduce the deviation from normality for egg production.

3.2 Materials and methods

3.2.1 Data description

The two populations shared a common origin from a female line. The British population included 20784 animals, while the North American population numbered 16231 birds.

Table 3.1: Descriptive statistics covering all the four generations for body weight at 14 (BW14), 19 (BW19) and 24 (BW24) weeks of age for males (M) and females (F) and egg production (EGG).

Trait		**No records**	**Mean**	**SD**	**CV (%)**
UK POPULATION					
BW14 (kg)	M	10487	11.5	2.5	9.9
	F	10297	8.5	1.6	8.8
BW19 (kg)	M	3670	17.6	3 .2	8.1
	F	6385	12.2	1.9	7.1
BW24 (kg)	M	2688	22.3	3.5	7.2
	F	6261	14.6	2.2	6.8
EGG		1728	82.0	23.2	28.3
USA POPULATION					
BW14 (kg)	M	7859	11.6	2.1	8.26
	F	8372	8.3	1.4	7.55
BW19 (kg)	M	3442	17.4	2.7	7.11
	F	5083	11.4	1.6	6.36
BW24 (kg)	M	2452	20.7	3.8	8.22
	F	5435	13.2	2.0	6.8
EGG		1772	82.4	24.5	29.7

The description of this line can be found in Chapter 2. No data were available from the common base population before the separation. For the the USA population the records for BW19 were not available for female birds for year 2002. The summary statistics are presented in Table 3.1.

The available traits were: body weight at 14 (BW14), 19 (BW19) and 24 (BW24) weeks of age and the total number of eggs laid (EGG) within 139 days of production. In both locations the line was selected for body weight and egg number, as well as for body conformation, with an emphasis on increasing breast muscle yield. However, no data on this trait were available for the study.

3.2.2 Dealing with deviation from normality for egg production trait

Egg production in poultry has been shown to exhibit a significant departure from the form of a normal distribution (Clayton, 1975; Nestor *et al.*, 1996). The consequence is that genetic parameters estimates are biased when using REML, since the underlying assumptions of normality, linearity of heritability and homogeneity of variance are not satisfied (Ibe and Hill, 1988). A two-step method was used in this study to reduce the non-normality of egg production data.

The first step was to identify and exclude outliers from the dataset. In order to set a threshold to distinguish outliers from acceptable data, an approach similar to that of Koerhuis (1996) was used. In brief, egg production was treated as an all-or-none trait. It was assumed that individuals with egg production under a certain threshold do not contribute to the total genetic variance and are "outliers" for environmental reasons, e.g. disease. Their egg production was set equal to zero, while records of birds with higher production were recoded as one. The heritability was estimated using the linear model for various thresholds. The threshold for outliers was taken to be the value for which heritability of the recoded data exceeded 0.001, a reasonable but nevertheless arbitrary threshold. Data on birds with production below the threshold were excluded from the main analysis, i.e. treated as missing data for all traits.

The second step was the application of the Box-Cox transformation Box and Cox (1964) to the egg production data. Previous research showed that this methodology reduces the heterogeneity of residuals, increases the linearity of heritability and provides a satisfactory approximation to the normal distribution for egg production (Ibe and Hill, 1988; Besbes *et al.*, 1993; Koerhuis, 1996; Savas *et al.*, 1999). The formula for this transformation is:

$$\frac{y^{\lambda}-1}{\lambda G_y^{\lambda-1}} \tag{3.1}$$

where y is the original untransformed observation and G_y is the geometric mean of the *y*'s.

In order to estimate the optimum parameter λ , several runs of REML models, described in the next section, with different values were performed to produce the log-likelihood profile for λ (Darwash *et al.*, 1997). The optimal λ value was chosen as that which maximised the log likelihood for the additive animal model (Meyer, 1989). However, rounded values were preferred, unless the final model likelihood was notably different in the log-likelihood tests from that of the intermediate optimum value. The rationale for this preference was that rounded values to the nearest integer (or half-integer) for λ are more easily interpreted.

3.2.3 Genetic model

Rearing and laying farms as well as hatches were found to be confounded and no data were available for certain hatches and years. Hence, a fixed effect, growth cohort, representing all year, rearing and hatch subclasses was fitted. Similarly, for egg production the egg cohort, representing all year, laying farm and pen subclasses, was used. Sex and its interaction with growth cohort were included as fixed effects in analyses of body weight. All data were analysed with the ASReml package (ver.1.10 Gilmour *et al.* (2001)).

In order to obtain unbiased estimates of variance components, a multivariate model was applied, including four traits: BW14, BW19, BW24 and EGG, since this would cater for the impact of consecutive selection. Several models of analysis were compared with the base model (model (O), no dam effect) and between them:

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z_{1i}a_i} + \mathbf{e_i} \tag{O}$$

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z_{1i}a_i} + \mathbf{Z_{2i}c_i} + \mathbf{e_i}, \quad cov(\mathbf{a_i}, \mathbf{c_i}) = 0 \tag{A}$$

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z_{1i}a_i} + \mathbf{Z_{2i}d_i} + \mathbf{e_i}, \quad cov(\mathbf{a_i}, \mathbf{d_i}) = 0 \tag{B}$$

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z_{1i}a_i} + \mathbf{Z_{2i}d_i} + \mathbf{e_i}, \quad cov(\mathbf{a_i}, \mathbf{d_i}) = \mathbf{A}\sigma^2 a_i d_i \tag{C}$$

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z_{1i}a_i} + \mathbf{Z_{2i}d_i} + \mathbf{Z_{2i}c_i} + \mathbf{e_i}, \quad cov(\mathbf{a_i}, \mathbf{d_i}) = 0 \tag{D}$$

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z_{1i}a_i} + \mathbf{Z_{2i}d_i} + \mathbf{Z_{2i}c_i} + \mathbf{e_i}, \quad cov(\mathbf{a_i}, \mathbf{d_i}) = \mathbf{A}\sigma^2 a_i d_i \tag{E}$$

where $\mathbf{y_i}$ (N_i) is the vector of N_i observations recorded for the i^{th} trait; $\beta_\mathbf{i}$ (f_i) is the vector of fixed effects for the i^{th} trait; $\mathbf{a_i}$ (N_i) is the vector of the random additive genetic effect for the i^{th} trait ($\mathbf{a_i} \sim N(0, \sigma^2 a_i)$); $\mathbf{c_i}$ (N_p) is the vector of the environmental permanent maternal effects common to all the progeny from a given dam; $\mathbf{d_i}$ (N_p) is the vector of maternal genetic effects ($\mathbf{d_i} \sim N(0, \sigma^2 d_i)$); $\mathbf{e_i}$ (N_i) is the vector of residuals for the i^{th} trait ($\mathbf{e_i} \sim N(0, \sigma^2 e_i)$); $\mathbf{X_i}$, $\mathbf{Z_{1i}}$, $\mathbf{Z_{2i}}$ are known design matrices; and $\mathbf{A}$ is the known relationship matrix between animals. Extension of previous notations to the multivariate analysis of $\mathbf{y_i}$ and $\mathbf{y_j}$ was straightforward with for (A), (D) and (E) cov($\mathbf{a_i}, \mathbf{c_j}$)=0 ($i \neq j$), for (B) cov($\mathbf{a_i}, \mathbf{d_j}$)=0, for (C) and (E) cov($\mathbf{a_i}, \mathbf{d_i}$)=$\mathbf{A}\sigma^2 a_i d_i$ ($i \neq j$).

The log-likelihood was computed for all five models and that used for the final analyses was chosen using the principle of Occams razor (Feurer, 1957).

3.3 Results

3.3.1 Outlier exclusion & Box-Cox transformations

The test applied for the outlier identification indicated that the optimum threshold excluded individuals whose egg production was equal to or less than 15 eggs in the UK population and 5 eggs in the USA. With these criteria, 69 and 53 birds respectively were excluded, representing approximately 4% and 3% of the total number of records. The exclusion of these individuals reduced the skewness and kurtosis (Table 3.2). After consecutive REML runs, the optimum values for parameter λ of the Box-Cox transformation were found to be 2.70 and 2.60 respectively for the UK and USA populations. However, the λ-parameters finally selected were 3 (i.e. a cubic transformation) and 2.5, since the reductions in the log-likelihood from the optimum using these values were not significant (P>0.05). After the application of Box-Cox transformation both the skewness and kurtosis were considerably reduced in magnitude, suggesting a better approximation to normal distribution (Table 3.2). This is confirmed in Figure 3.1, showing histograms for

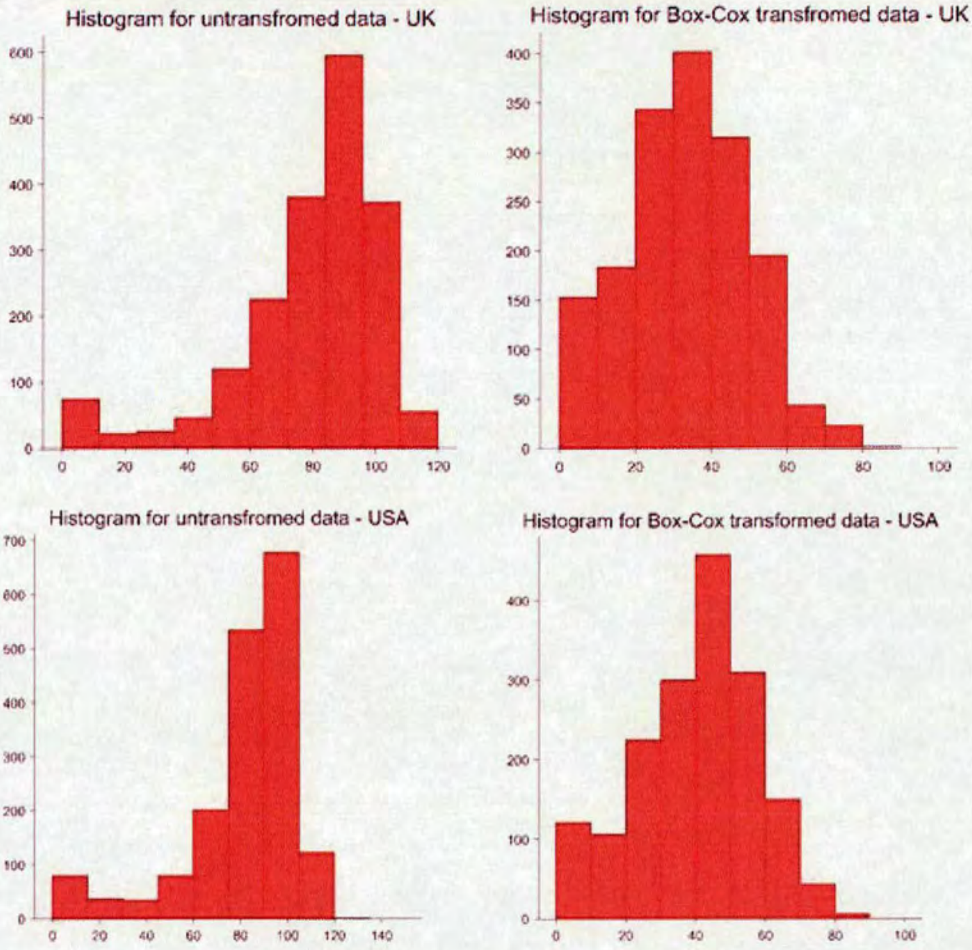
Table 3.2: Effects of outlier exclusion (step 1) followed by Box-Cox transformation (step 2) on the coefficient of variation (C.V.), skewness and kurtosis of egg production.

	Raw Data	**After Step 1**	**After Step 2**
UK POPULATION			
Number of observations	1728	1659	1659
C.V.	29.5	20.9	13.1
Skewness	-1.56	-1.26	0.03
Kurtosis	2.57	1.99	-0.25
USA POPULATION			
Number of observations	1772	1719	1719
C.V.	29.7	23.4	15.7
Skewness	-1.76	-1.26	-0.37
Kurtosis	2.97	1.99	-0.30

each population of the distribution of the egg production data when untransformed and after Box-Cox transformation (step 2). Figure 3.2 shows a normal probability plot for the UK population in order to assess visually whether or not the transofrmation reduced the deviation from the normal distribution. The corresponding graphs for the USA population were very similar (results are not shown).

3.3.2 Maternal effects model

In total one base model including only an animal effect and five maternal models were tested for both body weight traits and egg production (Table 3.3). The principal result was that the inclusion of a second random effect, representing maternal effects, either genetic or environmental, gave a significant increase in the log-likelihood for all the body weight traits. In contrast, maternal effects appear not to influence the egg production and the maternal random effect was omitted from all subsequent analyses reported. For body



Histogram for untransfromed data - UK

Histogram for Box-Cox transfromed data - UK

Histogram for untransfromed data - USA

Histogram for Box-Cox transformed data - USA

Figure 3.1: Histograms for the untransformed (left panel) and the Box-Cox transformed (λ=3 for UK & λ=2.5 for USA) data (right panel) for total egg number in selection flocks of turkeys in the UK and USA.

weight, there was no evidence ($P>0.05$) supporting the inclusion of both an environmental and a genetic maternal effect (comparing models D, E to A, B). Also in the case of the genetic model, there was no evidence of covariance between additive and maternal effects (i.e. model C). Therefore the choice of the final model was made between A and B. Model A was chosen, because it introduces no covariances between maternal effects of different individuals.

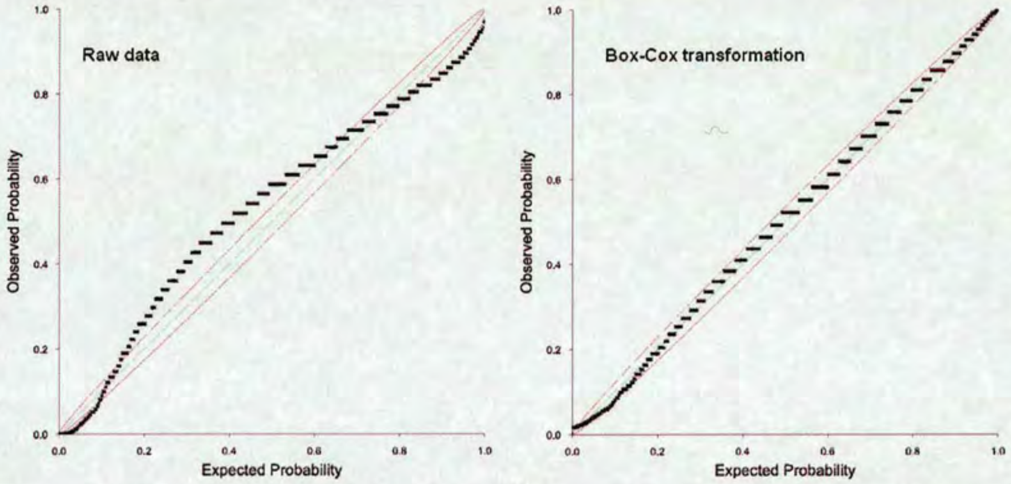


Figure 3.2: Normal probability plots (observed probability - expected probability) for the untransformed (left panel) and the Box-Cox transformed (right panel) in the UK population.

3.3.3 Multi-trait analysis

The estimates of the variance components are presented in Table 3.4. The additive genetic variance (σ_A^2) for the USA birds was higher. For body weight, the common environmental variance (σ_c^2) only represented approximately 4% of the phenotypic variance (σ_p^2) for UK and 2% for USA populations (using model A as described in the previous section). Whilst the fraction of the phenotypic variance of body weight explained by the maternal effects was low, it was statistically significant for all body weight traits, as the log-likelihood tests suggested. Finally, the residual variance (σ_e^2) was lower for the USA birds.

The genetic and phenotypic parameter estimates are shown in Table 3.5. In general, the precision of the heritability estimates was higher for body weight traits than that for egg production, since fewer records were available for the latter. Heritability estimates for UK birds were lower than USA birds. Genetic correlations between weight measures taken between 14, 19 and 24 weeks ranged between 0.89 and 0.97 in the two populations. In contrast, the genetic correlation between egg production and body weight was negative,

Table 3.3: Log-likelihood values for the model with only additive effects and their difference for all maternal effects models tested for body weight traits. All models differed significantly from the simple model, but no difference was detected between them *(A = maternal effects only as environmental ones; B = maternal effects only as genetic ones; C = same as B but with covariance between additive animal and maternal genetic effects; D = maternal effects as genetic and environmental; E = same as D but with covariance between additive animal and maternal genetic effects).*

Model[1]		**BW14**	**BW19**	**BW24**
Simple	UK	-19395	-10402	-11232
	USA	-14672	-8535.1	-8781
(A)	UK	24.1	4.1	0.6
	USA	10	3	1.9
(B)	UK	24.2	5	0.6
	USA	11.1	3.3	2.9
(C)	UK	24.1	5	0.2
	USA	11.5	3.5	2.9
(D)	UK	22.5	4.3	0.2
	USA	11.9	2.6	2.3
(E)	UK	22.4	4.2	0.5
	USA	7.5	2.8	2.5

strong and consistent over the different ages. However the magnitude of this association appeared to differ between the two populations: for UK birds the genetic correlation reached a value of -0.75, while for the USA population it did not exceed -0.55. These estimates were not statistically different, as it can be inferred from comparing the respective standard errors from Table—3.5. Estimates for the residual correlations (-0.06 to -0.20) were lower than those of the phenotypic ones and followed the correlation trends between the traits, highlighting the large difference between the gene action and the environmental effect for this particular line. The phenotypic correlations between body weight at dif-

Table 3.4: Estimates from multivariate REML of variance components for body weight at 14 (BW14, kg^2), 19 (BW19, kg^2) and 24 (BW24, kg^2) weeks of age and for total egg number laid in a production period of 139 days (EGG). Maternal effects were not fitted for egg number *(σ_A^2 = additive genetic variance; σ_c^2 = common environmental variance; σ_e^2 = residual variance; σ_p^2 = phenotypic variance)*

	BW14	**BW19**	**BW24**	**EGG**
UK POPULATION				
σ_A^2	1.052	1.762	2.046	1.145
σ_c^2	0.126	0.211	0.226	-
σ_e^2	1.688	3.285	5.017	4.133
σ_p^2	2.866	5.258	7.289	5.278
USA POPULATION				
σ_A^2	1.325	1.998	2.543	2.251
σ_c^2	0.062	0.087	0.123	-
σ_e^2	1.376	2.546	3.303	4.461
σ_p^2	2.763	4.631	5.969	6.712

ferent ages and egg production were also negative, but less strong, ranging from -0.14 to -0.26 for the UK population and around -0.15 in the USA.

3.4 Discussion

A strong negative genetic correlation between body weight and total egg production was detected in this study and its magnitude indicates that the simultaneous selection for both traits may not be effective. An on-going selection of alleles with pleiotropic but antagonistic effect could be an interpretation of these results.

The outlier exclusion and especially the Box-Cox transformation resulted in an approximation to the normal distribution. For both datasets, the optimal value of the power index

Table 3.5: Phenotypic correlations and genetic parameter estimates from the multivariate REML for body weight at 14 (BW14), 19 (BW19) & 24 (BW24) weeks of age and egg production (EGG). Heritability estimates are given on the diagonal (in bold), genetic correlations above diagonal and phenotypic below.

	BW14	**BW19**	**BW24**	**EGG**
UK POPULATION				
BW14	**0.37 ± 0.03**	0.96 ± 0.01	0.89 ± 0.02	-0.74±0.07
BW19	0.65	**0.34 ± 0.03**	0.96 ± 0.01	-0.75±0.08
BW24	0.53	0.79	**0.28 ± 0.03**	-0.75±0.08
EGG	-0.14	-0.22	-0.26	**0.22±0.04**
USA POPULATION				
BW14	**0.48 ± 0.03**	0.97 ± 0.01	0.89 ± 0.01	-0.48±0.09
BW19	0.72	**0.43 ± 0.03**	0.97 ± 0.01	-0.51±0.08
BW24	0.69	0.77	**0.43 +0.03**	-0.55±0.09
EGG	-0.15	-0.15	-0.14	**0.34±0.06**

for transformation (λ) was very similar and greater than 1 as expected, since raw egg production data is negatively skewed (Ibe and Hill, 1988). However, because λ relies heavily on the specific dataset structure, the values proposed by previous researchers cannot be compared directly, especially since they referred to chicken populations (Ibe and Hill, 1988; Besbes *et al.*, 1993; Koerhuis, 1996; Savas *et al.*, 1999; Unver *et al.*, 2004). Nevertheless, the power transformations of 2.6 and 2.7 obtained for these data were remarkably similar to those of Chapuis *et al.* (1996), who reported values of 2.75 and 2.4 respectively, although the turkey populations for the two studies were different.

The application of the two-step procedure had an impact in increasing heritability. The estimated heritability for the untransformed data in the UK population for the egg production was 0.11, slightly increased after step 1 to 0.13, while after the Box-Cox transformation (step 2) it reached 0.22. However, the first step removed the observed bi-modality in such a way as to limit the loss of genetic variation. Similar results were obtained for the

USA population.

It was concluded that the inclusion of maternal effects had an impact on modelling body weight but not for egg production. Ignoring maternal effects resulted in the overestimation of the heritability for body weight, a conclusion also reached by others (Chambers, 1990; LeBihan-Duval *et al.*, 1998). However, the form of the maternal effects fitted has differed among studies; Koerhuis (1996) and LeBihan-Duval *et al.* (1998) treated the maternal influence in broilers as an environmental effect, whereas Chapuis *et al.* (1996) fitted genetic effects in their analysis of turkey data. Furthermore, Koerhuis *et al.* (1997) proposed a model including both additive genetic and environmental maternal effects when studying juvenile body weight at 6 weeks of age. Nevertheless, the dataset used in this study included body weight measurements at a greater age, when maternal influences could be assumed to have less input. Our results supported the hypothesis that the contribution of maternal influences to the phenotypic variance decreases with age (Table 3.3). Furthermore, the lack of evidence for maternal effects for egg production is also consistent with this hypothesis. The failure to detect maternal effects in egg production is common to other poultry studies (see Szwaczkowski (2003) for a review).

In the population structure studied here, there dams are nested within sires, so there are mostly full sibs (with few exceptions of remates) and all chicks were hatched within a single laying season. Thus, it is difficult to separate the dominance variance from environmental covariances among sibs, or other causes of maternal effects, e.g. egg-related factors. The low contribution of the maternal effects to the total variance suggests that the dominance effect within these lines is low: the estimated value of $\sigma_c^2 = 0.04$ places an upper bound of 0.16 on the fraction of variance explained by dominance ($\sigma_d^2 \sim 4 \times \sigma_c^2$). However, this does not imply that inbreeding depression would be absent since the magnitude of the depression will depend on other factors such as the number of loci contributing to this variance and the degree of directional dominance. So, it does not preclude a cross with another line may display significant heterosis. The latter is an important observation given the structure and the potential options available to the turkey breeding industry.

In this study body weight was not treated as a separate trait in males and females, in contrast with previous suggestion from Chapuis *et al.* (1996). In an analysis where sexes were analysed separately (results not shown), heritability estimates for body weight were marginally higher for females, consistent with the observations of LeBihan-Duval *et al.* (1998). However, the genetic correlation between male and female body weight was very high (0.95) and the genetic correlation estimates between body weight and egg number were similar when the analysis included only female records. This justified the use of body weight as a common trait for both sexes to reduce dimensionality, especially since no information for an underlying genetic mechanism is available that would suggest the opposite. Nevertheless, sex was used as a fixed effect since its effect on body weight was significant.

The heritability estimates for both body weight and total egg number were within the range suggested in the literature. Arthur and Abplanalp (1975) reviewed 18 reports on the heritability of body weight at various ages and arrived at an average of 0.41 and Buss (1990) reported a similar average but indicated that estimates considered by him to be reliable could range from 0.23 to 0.71. For total egg number, Arthur and Abplanalp (1975) reported a range from 0.13 to 0.49 with an average of 0.22. The most relevant comparison of parameters is with Chapuis *et al.* (1996) who used large samples from two modern female lines, in which combined selection for growth and egg production was performed, with a similar transformation to egg number, in order to estimate genetic parameters. Their estimates for body weight were very high, exceeding 0.50 for body weight at 12 and 16 weeks, a comparable value to the result for the USA flock, while they reported heritability estimates for egg production of over 0.30 for Box-Cox transformed data.

In the results, the heritability estimates of the USA population were greater than that for the UK population although they originated from the same base population. This was due to both a greater genetic variance for all traits, and, for body weight, less environmental variance. There is a number of possible explanations: different environments, e.g. climate and dietary ingredients encouraging greater expression of genetic potential; or genetic sampling in establishing the lines. However, the lack of data from the common generation prevented a detailed study of the genotype-environment interactions.

Whilst the genetic correlations among the body weights were all close to 1, giving flexibility to select for body weight as early as 14 weeks of age, the genetic correlations observed between body weight and egg production were large in magnitude and antagonistic. The values for the UK and USA populations averaged -0.72 (s.e. 0.08) and -0.50 (s.e. 0.09) respectively, among the largest reported observations for turkeys. For female lines, Arthur and Abplanalp (1975) and Buss (1990) reported much lower consensus values of -0.1 and -0.05 respectively, yet these results refer to lightweight strains where selection pressure for growth was minimal.

In heavier populations, Nestor (1977a) reported realised genetic correlations of -0.42 (s.e. 0.12) for a line selected for body weight, and a more negative -0.66 (s.e. 0.42) for a line selected for growth and egg number, but with a correspondingly large standard error. Further evidence is provided by Chapuis *et al.* (1996) where their estimates ranged from -0.2 (light line) to -0.5 (heavy line). The genetic correlation was even stronger for the populations included in this study being consistent with the fact that the birds were heavier.

Hence, it may be suggested that the magnitude of the correlation for these modern lines with heavier body weight is much greater than proposed by Arthur and Abplanalp (1975) and Buss (1990) and a more reasonable consensus value might be -0.53 (a weighted average from the present study, Chapuis *et al.* (1996), Nestor *et al.* (1996) and Emmerson *et al.* (2002). Such a consensus value offers a very different prospect for selection towards increasing body weight and maintaining egg production.

Furthermore, Nestor *et al.* (1996) gave results of long-term response of egg production to continued selection solely for increased body weight and showed that the rate of response declined over time. Starting from a lower average egg production than any of the lines studied here, the egg production decreased and reached an apparent selection plateau, where perhaps further decreases were offset by natural selection. This could arise from fixation during selection of genes favouring larger weights but with minimal effect on egg production, leaving the segregating genetic variation dominated by pleiotropic loci with antagonistic effects on the traits. It also suggests that intense selection pressure for body weight could lead modern heavy female lines to such an unfavourable low egg production, despite the simultaneous selection for the trait, due to the lower accuracy of selection for egg production compared to body weight. In conclusion, the results presented in this analysis provided evidence to support the hypothesis as this was conceptualised in Figure 1.3.

One other indirect consequence of the intense selection for growth seems to be the more erratic laying patterns in the populations studied. Traits that take account of the number of the eggs set in a specific period, such as the persistency and intensity of lay, might provide further information to increase the accuracy of selection for higher egg numbers. For the latter, evidence exists that it mediates the strength of the correlation between body weight and egg number (Nestor *et al.*, 2000). Such a hypothesis may warrant an investigation into alternative selection strategies for heavy turkey lines, including a greater role for egg laying pattern traits.

Within this context, the next chapters provide a longitudinal analysis of the egg production in turkeys. The next chapter explores the potential use of the intensity of lay, defined as a time-to-event trait, while Chapters 5, 6 and 7 focus also on the persistency of lay, expressed as individual laying patterns and how these can be capitalised on the more detailed genetic analysis of the egg production.

Chapter 4

Application of Survival Analysis for the genetic study of time-related patterns of egg production

4.1 Introduction

Survival analysis involves the study of the time between a fixed starting point and a terminating event that may occur within a specified period. It is possible that for some subjects included in the analysis no event occurs, or these records are lost from the follow-up. This phenomenon is known as right censoring and requires special methods of analysis (Cox, 1972). The advantage of survival analysis is that it combines the information on whether records are censored or not in order to model effectively time-to-event traits.

An early study in animal breeding employing survival analysis for the statistical analysis of mortality using non-simulated data dates back to 1986 (Woolliams *et al.*, 1986). The potential of this methodology for genetic evaluation was demonstrated in studies of low-heritable traits for which analysis is not straightforward, because their distribution is unknown or skewed (Sölkner and Ducrocq, 1999). Longevity constitutes a typical trait for which survival analysis has been applied to dairy cattle (Ducrocq, 1987; Vukasinovic *et al.*, 1999), lambs (Southey *et al.*, 2000), rabbits (Piles *et al.*, 2006), swine (Yazdi *et al.*, 2000), and poultry (Ducrocq *et al.*, 2000).

Recently, survival analysis has been applied to the genetic analysis of other time-related traits, such as the study of the time to the first veterinary treatment of clinical mastitis (Saebø and Frigessi, 2004) and the interval between calving and the last insemination (Schneider *et al.*, 2005) in dairy cattle. In the former example it was suggested that the formulation of time-to-event traits may contribute in the more efficient ranking of sires by providing the probability of them being in the genetically superior group. Schneider (2006) showed that the use of survival analysis compared to a linear model for the longitudinal analysis of female fertility and mastitis traits can achieve higher accuracy of selection which translates to more rapid genetic progress.

The most common approach for the genetic analysis of egg production is to use the total number of eggs laid within a specified period. However, as discussed in Chapter 3, the total egg number is a discrete character and therefore, it is characterised by a significant deviation from normality that hinders the estimation of genetic parameters. Although power transformations to data can be applied to reduce this departure from the normal distribution, it may also be possible to use an alternative approach. The survival analysis framework can provide such a tool.

In order to formulate a time-to-event trait to use in a survival analysis, two points have to be defined for time recording: a starting and a terminating point. In the context of the study of egg production the beginning of the time of recording can be taken as with the onset of the laying period, while the "event" can be defined as the days that are required for a hen to lay a specific number of eggs. If a hen fails to reach the production target or it is culled before this happens, then it is treated as a censored record. The survival trait is complementary to the cumulative egg production over a specific period in that the former measure accounts for a fixed number of eggs laid over time, while the latter accounts for the eggs laid over a fixed interval. Both measures capture features of laying dynamics, but under the survival framework setting a starting point and a production target allow focusing on specific laying stages and studying the performance of birds over time.

The aim of this chapter was to investigate the application of survival analysis to the longitudinal genetic analysis of egg production. This was an exploratory study, as this has not been applied to poultry before. Therefore, non-parametric models were initially fitted for determining an appropriate parametric model to permit the inclusion of a random sire effect. The aim was to obtain heritability estimates to quantify the contribution of the genetic component on laying patterns.

4.2 Materials and methods

4.2.1 Key concepts of survival analysis

Let T be a variable for the time to an event for an individual. Then the probability that T exceeds a specified time t is given by the survival function $S(t)$:

$$S(t) = P(T > t) = 1 - F(t) \tag{4.1}$$

where $F(t)$ is the cumulative distribution function for T. Therefore, the survival function represents the probability that the event does not occur to an individual from the starting point to a specified time. In the context of this analysis the survival function provided the probability that a hen has not laid a pre-specified number of eggs up to a specific day of the laying period. In order to model the survival function, a hazard function can be defined. This function represents the instantaneous risk that the event occurs to an individual in the interval Δt conditional upon survival up to time t. So, the hazard function $\lambda(t)$ is given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{4.2}$$

where $f(t)$ is the probability density function of T expressing the change in the survival function within a small time interval Δt (i.e. $f(t) = S'(t) = \frac{dS(t)}{dt}$). The relationship between the survival and hazard functions can be derived from equation 4.2 as:

$$S(t) = exp[- \int_0^t \lambda(u) du.] \tag{4.3}$$

$$\lambda(t) = -\frac{S'(t)}{S(t)} = \frac{d}{dt}(-logS(t)) \tag{4.4}$$

The survival function can be estimated directly from the data using a non-parametric model. Alternatively, the hazard function can be defined using a semi or fully parametric model and as follows from 4.3 and 4.4 the corresponding survival function can be estimated. In this study both non and fully parametric models were considered. These are presented in the following sections.

4.2.2 Nonparametric survival model

Empirical estimates of survival and hazard functions were obtained using the nonparametric Kaplan-Meier (KM) or product-limit method (Kaplan and Meier, 1958). A general formula for obtaining KM survival probabilities is:

$$\hat{S}(t_j) = \hat{S}(t_{j-1})(1 - \frac{d_j}{n_j}) \tag{4.5}$$

According to 4.5, the KM estimate for an individual of not having the event at time j is calculated from the survival probability at time t_{j-1}, S_{j-1}; the number of individuals not having the event just before time t_j, d_j; and the number of events occurring at time t, n_j. A fundamental assumption of the KM method is that events occur independently and thus it is possible to multiply the survival probabilities together in order to obtain the cumulative survival probability. KM survival probabilities can be plotted against time as a step function for each factor affecting survival, providing a quick and robust visualisation of their impact on the time-to-event. KM curves for each group can be compared to test if they are statistically equivalent or not by using a logrank test (Peto and Peto, 1972) that employs a Chi-square test to compare the expected and observed number of events in each group.

In this study KM curves were plotted for the factors "year" in order to detect trends that may be due to the on-going selection for body weight and for "sire" in order to investigate whether there was a significant variance between groups of hens stratified as paternal half-sibs. The corresponding KM curves were compared using the logrank test. This evidence was used to develop more sophisticated models, as described below. The nonparametric analysis was carried out using SPSS and STATA packages.

4.2.3 Parametric survival model

When using nonparamteric survival curves it is not possible to consider more than one factor simultaneously (Bradburn *et al.*, 2003). Moreover, a logrank test can only indicate a difference between groups, but does not provide an estimate of the effect. A popular method to overcome these constraints is to use a regression model for the hazard function. Thus, the hazard of an individual can be estimated at any time, given specification of a set of explanatory variables, according to the following formula

$$\lambda(t,\mathbf{X}) = \lambda_0(t)exp\{\sum_{i=1}^{p} \beta_i \mathbf{X}_i\} \tag{4.6}$$

where λ_0 is the baseline function and $\mathbf{X}_i$ is the vector of predictor variables with the associated regression coefficients β. The first component of the model, the baseline hazard, is a function of time and the second, the exponential expression, involves only the effects X's, which do not depend on time. Equation 4.6 encapsulates the proportional hazard assumption (PH), according to which the hazard of one individual is proportional to the hazard of any other individual independently of time, or in other words that the hazard ratio remains constant over time. The PH assumption in mathematical notation translates as follows:

$$\frac{\hat{\lambda}(t,\mathbf{X})}{\hat{\lambda}(t,\mathbf{X}^*)} = \hat{\theta} \tag{4.7}$$

Equation 4.7 shows that when the hazard functions of any two individuals are plotted against time, the curves cannot cross. The most general regression model that satisfies the PH assumption is the Cox model (Cox, 1972). According to this, the baseline function is not defined and this is why this model is also called a semi-parametric model. Although the baseline hazard is unspecified, reasonably good estimates of the regression coefficients can be obtained (Kleinbaum and Klein, 2005). The PH assumption can be tested graphically by comparing the expected survival curve derived from a Cox model against the observed KM. The PH assumption is valid only when the estimates and observed curves are close to each other.

In contrast to the Cox model the form of the baseline hazard function is completely specified in a parametric model. The Weibull model is a widely used parametric model in which the survival times are assumed to follow the Weibull distribution. The Weibull model satisfies the PH assumption, but in addition it also implies that the effect of explanatory variables on the hazard is monotonic with survival time. The hazard function of a Weibull model is defined as follows:

$$\lambda(t,\mathbf{X}) = \gamma \rho t^{\rho-1} \tag{4.8}$$

$$where\, \gamma = exp\{\sum_{i=1}^{p} \beta_i \mathbf{X}_i\}\,,\, \gamma > 0$$

$$and\, \lambda_0(t) = \rho t^{\rho-1}$$

The parameter γ is derived from the regression coefficients. The parameter ρ is called shape parameter and it determines if the hazard increases ($\rho > 1$), remains constant ($\rho = 1$, Weibull reduces to the exponential model) or decreases ($\rho < 1$) over time. The baseline hazard λ_0 corresponds to the intercept of the model.

Preliminary results from the empirical hazard function suggested that the hazard increases monotonically over time and it was hypothesised that a Weibull model might fit the data. In order to evaluate this hypothesis, log-log KM survival curves were plotted against the logs of time. If a linear relationship is to be detected on the graphical test, this suggests that a Weibull model is suitable to fit data on a fully parametric model. The benefit from fitting a Weibull model to data is a significant reduction of the computation time compared to the Cox model, particularly for a large sets of records (Sölkner and Ducrocq, 1999). Except for specifying the baseline hazard function, it was also important to determine which factors will be included in the model. Log-likelihood tests were performed comparing the likelihoods obtained from the candidate Weibull models using the STATA package.

4.2.4 Frailty survival model

The set of the explanatory variables included in the Weibull model corresponds to the fixed effects in a linear models context. It is possible to extend the model to include also a random effect. These models are called frailty models (Vaupel *et al.*, 1979). An extension is the shared frailty model, in which groups of individuals are assumed to share the same frailty (Kleinbaum and Klein, 2005). In the context of a genetic analysis, this similarity refers to the shared features between half/full sibs due to both genetic and environmental factors.

In this study the Weibull model with two fixed effects, year and pen, was extended to include a random effect accounting for the sire component. Thus, the hazard function for time t of hen i was defined as follows:

$$\lambda_{ijkm}(t) = \lambda_0(t) e^{year_j + pen_k + sire_m} \tag{4.9}$$

where λ_0(t) is the baseline hazard function assumed to follow a Weibull distribution with parameters γ and ρ as defined in 4.8; $year_j$ the fixed effect of year *j*; pen_k the fixed effect of pen *k*; and $sire_m$ the random effect of sire *m*, the sire of hen *i*. Sire effects were assumed to follow a multinormal distribution with variance σ_s^2. Sires were assumed to be unrelated.

4.2.5 Estimation of genetic parameters

The sire model presented in Equation 4.9 does not include any residual component. Various approaches have been suggested to account for the residual variance. In this study, heritability estimates were obtained on the original scale, using the equation proposed by Yazdi *et al.* (2002), which is not dependant on the Weibull parameters provided that parameter ρ does not take extreme values. The authors showed that when ρ takes values around 2, the heritability estimates from the following simple equation are close to estimates derived from more complex formulas (Ducrocq, 1987):

$$h^2 = \frac{4\sigma_s^2}{\sigma_s^2 + 1} \tag{4.10}$$

where σ_s^2 is the sire variance, which was estimated using the Survival Kit v3.12 (Ducrocq and Sölkner, 1998). The error of the heritability estimate was computed using the formula from Gilmour *et al.* (2001):

$$var(\frac{\sigma_a^2}{\sigma_p^2}) = (\frac{\sigma_a^2}{\sigma_p^2})^2(\frac{Var(\sigma_a^2)}{\sigma_a^4} + \frac{Var(\sigma_p^2)}{\sigma_p^4} - \frac{2Covar(\sigma_a^2,\sigma_p^2)}{\sigma_a^2\sigma_p^2}) \tag{4.11}$$

Given that $\sigma_a^2 = 4\sigma_s^2$, $\sigma_p^2 = 1 + \sigma_s^2$, the variance of the heritability is given by the following formula:

$$var(\frac{\sigma_a^2}{\sigma_p^2}) = \frac{16\sigma_s^4(\frac{Var(\sigma_s^2)}{\sigma_s^4} + \frac{Var(\sigma_s^2)}{(1+\sigma_s^2)^2} - \frac{2Var(\sigma_s^2)}{\sigma_s^2(1+\sigma_s^2)})}{(1+\sigma_s^2)^2} \tag{4.12}$$

The standard error of the heritability was computed as the square root of the estimate in 4.12.

4.2.6 Defining starting point and terminating event

The dataset consisted of daily egg production records that covered a period of 140 days. The average yield across all the five available generations was 82 eggs. This was considered as the target production and hence, the "event" was defined as the time that a hen required to lay 82 eggs.

The definition of the event affected the selection of the starting point for measuring the time to event. Due to the biological limitation that a hen can only lay one egg per day, it is implied that at least 82 days are required before the event occurs. In reality more time is needed, because eggs are laid in clutches. Although the intuitive choice was the onset of lay, preliminary experimentation suggested that using day 90, which coincided with the day that the first hen reached the target production, gave better results, in terms of fitting the model. The rationale of this choice is discussed in the following sections.

In conclusion, the dataset consisted of 2289 individuals, from which 1599 hens were uncensored (they laid 82 eggs before the 140 days) and 688 were censored (their production was less than 82 eggs over 140 days). The average time-to-82-eggs was 30 days after the starting point (which corresponded to day 120 of the whole laying period).

4.3 Results

4.3.1 Nonparametric analysis (KM curves)

A plot of KM estimates for the survival and hazard function stratified per year is presented in Figure 4.1. The trend observed from this graph implies that the rate of lay was reducing across years. The results from the logrank test suggested that there was sufficient evidence to reject the null hypotheis that the KM curves stratified per year were statistically equivalent ($p < 0.05$). This also implied that the effect of the year was significant and thus it should be included in the parametric model. The graph of the empirical hazard function

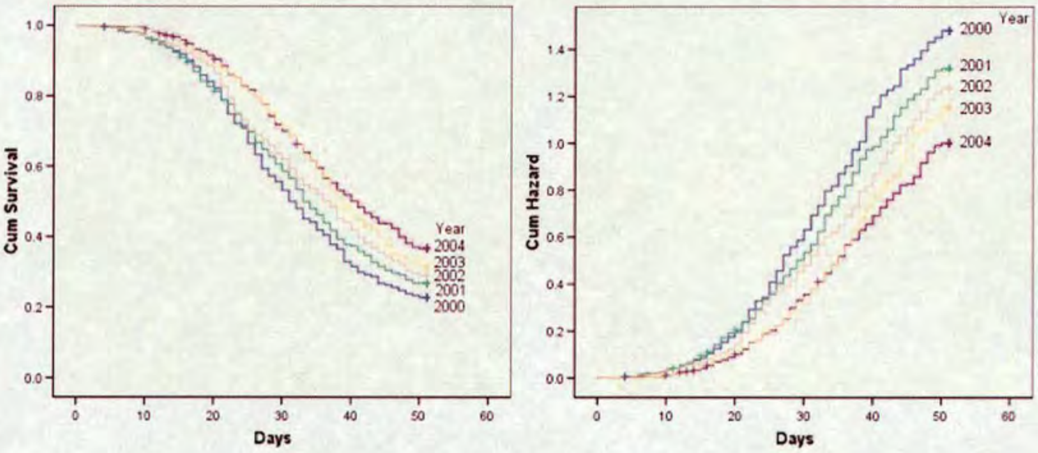


Figure 4.1: Kaplan-Meir plots of empirical survival (left panel) and hazard function (right panel) stratified per year.

indicated that hazard appeared to be monotonically increased over time.

A plot of the KM estimates for the survival and hazard function stratified per sire is presented in Figure 4.2. In the graph it is shown that a considerable variability is observed between KM curves corresponding to different sires. For certain sires all their progeny reached the production target, while for other sires, few of their offspring were successful. The results from the logrank test suggested that the KM curves were not equivalent for all sires ($p < 0.05$). This indicates that the sire component explained a significant part of the variance observed in the "survival" probabilities. The graph of the empirical hazard function also implied that hazard appeared to be monotonically increasing over time.

Similar KM estimates for the survival and hazard function stratified per pen were obtained, suggesting that pen also constituted a significant factor to be included in the analysis. The evidence accumulated from the nonparametric models was used to develop the fully parametric model. The results are presented below.

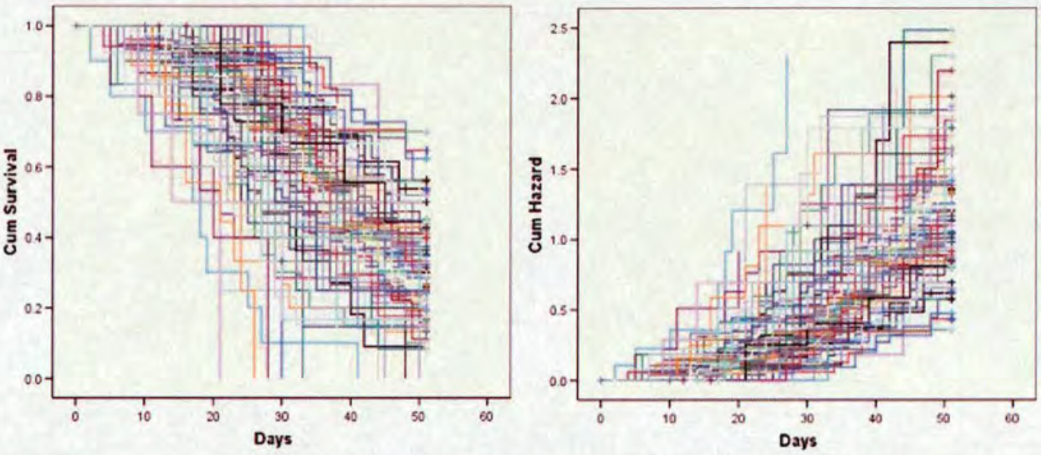


Figure 4.2: Kaplan-Meier plots of empirical survival (left panel) and hazard function (right panel) stratified per sire.

4.3.2 Parametric model (Weibull model)

It was concluded from the log-likelihood tests that two fixed effects had to be included in the Weibull model: year and pen. The inclusion of an interaction term was not found significant and therefore it was omitted.

In order to develop a parametric model, the proportional hazards assumption was graphically tested. In Figure 4.3 a comparison of the survival curves estimated from a Cox model and the observed survival curves both stratified by year is presented. As the graph suggests, the curves appear to be close to one another and without considerable discrepancies. This lends support that the PH assumption is not violated.

The next graphical test was to assess whether a Weibull model provided a suitable fit to the dataset. Figure 4.4 presents the plots of the log(-log) transformed KM "survival" estimates against the log of time when the time of recording started at the onset of lay (left panel) and at the 90^{th} day of the production period (right panel). As it is demonstrated in

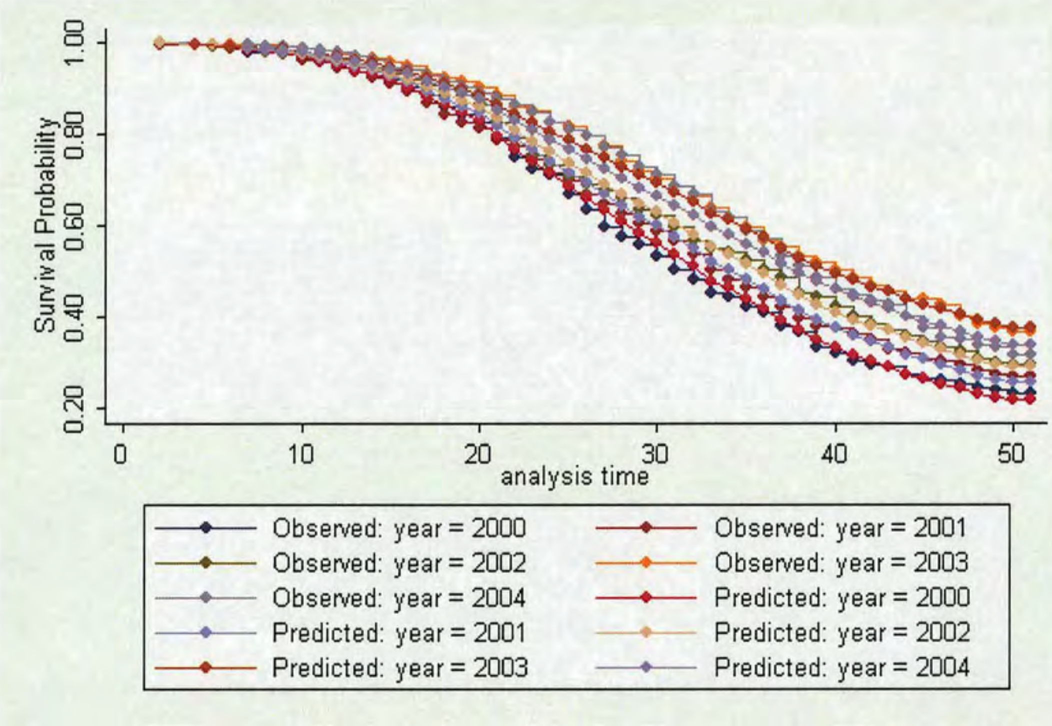


Figure 4.3: Comparison of survival curves expected from a Cox model and observed stratified per year against time. Lack of discrepancies between the curves suggests that the PH assumption is not violated.

the figures, only in the second case does a linear relationship appears to apply. Therefore, it can be concluded from this graphical test that the Weibull assumption is likely to hold only in the case in which time was recorded from the day that the first hen reached the production target (day 90).

The parameter ρ that determines the shape of the hazard function under the Weibull model was equal to 2.263. Since a value greater than one was found, this indicates that hazard increases over time, being consistent with the KM estimates. The intercept of the Weibull model λ_0 was estimated to be equal to -8.192. All regression coefficients from the Weibull model were significantly different from zero, providing evidence that all factors included in the model had a significant effect.

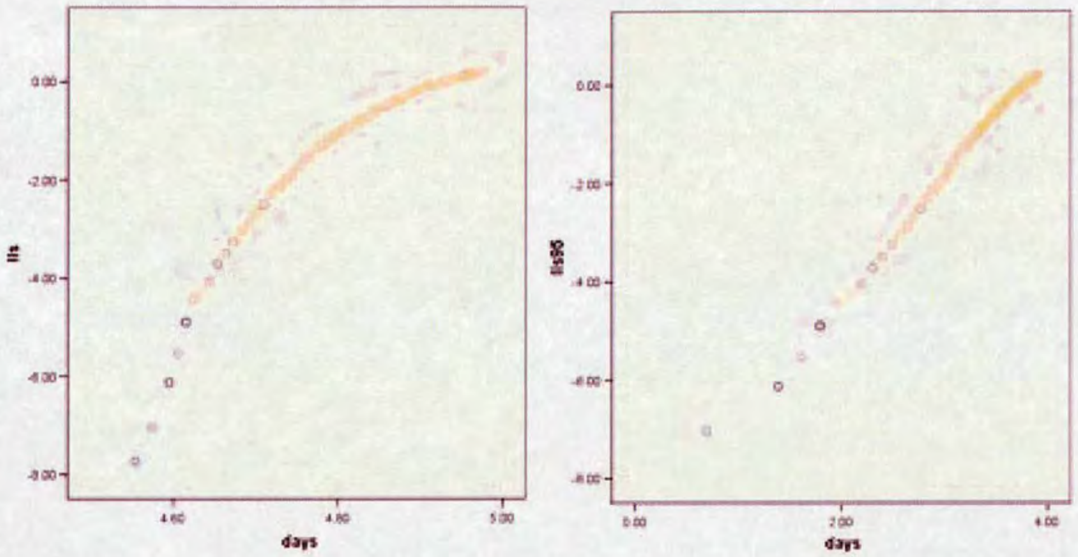


Figure 4.4: Log-Log Kaplan-Meier survival estimates stratified for year to compare the appropriateness of datasets in which time recording started the first (left panel and 90^{th} (right panel) day of the laying period.

4.3.3 Heritability estimates

The sire variance was estimated from the mode of the marginal posterior density and it was found to be equal to 0.047. The standard deviation of the marginal posterior distribution of the sire variance estimate was equal to 0.022. As reported above, the value of the Weibull shape parameter ρ was close to two, so it was possible to use the simplified heritability formula in Equation 4.10 in order to estimate the heritability of the time-to-lay-82-eggs. Hence, the point estimate of the heritability was 0.16, having a standard error equal to 0.08.

4.4 Discussion

This study constituted a preliminary investigation focusing on the time required to achieve a target egg yield, rather than the total production *per se*. The definition of the terminating event allowed me to encapsulate features of laying patterns, such as the persistency of lay, whilst using a Weibull model for modelling the baseline hazard. This simplified the estimation of the heritability, which was found to be similar to previous heritability estimates for total egg number in the same population. This study showed that the application of survival analysis is feasible and has potential to be a viable alternative for increasing the efficiency of selection for egg production.

As the breeding objective in the current turkey line was to increase egg production, the time-to-event trait should have a high correlation with the total egg number. In this analysis the event has been defined as the time required for hens to lay 82 eggs that corresponded to the average production over the five available generations. The phenotypic correlation between the time trait and the total yield was strong (0.71).

The time-to-event trait, as defined in this study, facilitated the distinction between fast and slow laying hens. Birds that reached the production target quickly, tended to lay more eggs in total than birds that required more days. In Figure 4.5 it is demonstrated that a strong linear relationship is observed between the rate of lay and the total production, when censored records are not considered. From the graph it is seen that, provided a hen achieves the production target, the time that it was required to the event is a good predictor of the total egg production. In conclusion, the definition of the terminating event allowed the formulation of a trait which was not only tightly linked to the total egg production, but also accounted for features of laying patterns, such is the rate and persistency of lay. The latter is of great importance for poultry breeders, since it is not easily defined (Grossman *et al.*, 2000), but is included in the breeding objectives (Groen, 2003).

The choice of the time event had also a positive impact on the shape of the hazard function. Since the number of hens reaching the target production increases over time, this implies that the probability that an event would occur in a short interval also increases towards

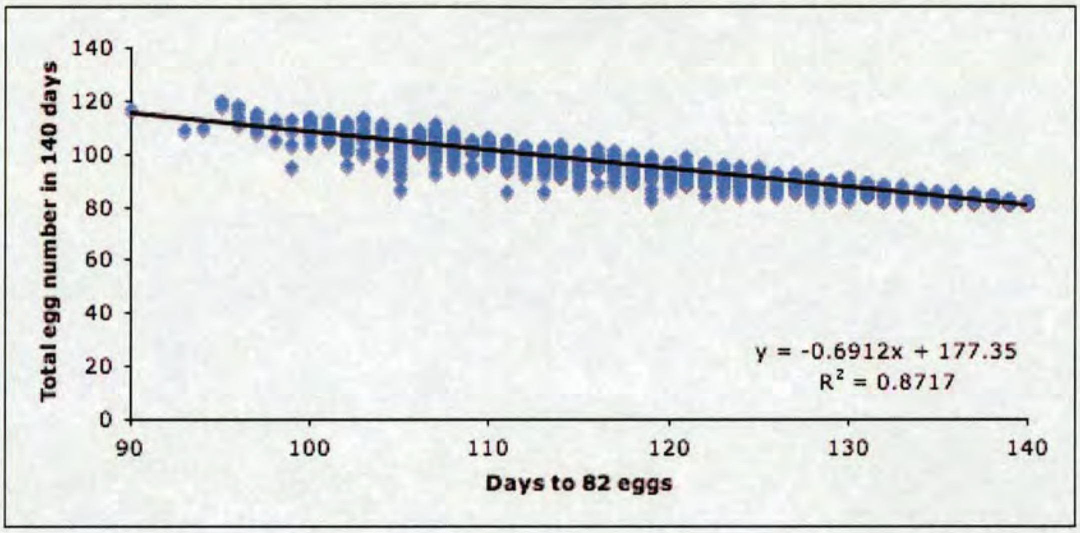


Figure 4.5: Fast laying hens tend to have higher total egg production, implying that the time-to-event trait may provide a means to assess persistency of lay (Censored records were not included in the graph.

the end of the laying period. This was consistent with the increasing hazards assumption and thus it was possible to fit a Weibull model with positive value for parameter ρ. Had the target production been set to correspond to a low number of eggs, the hazard would initially be increased, but after a period it would decrease, since almost all hens would have laid the required number of eggs. This would violate the assumption of monotonic hazard function and thus it would hamper the fit of a Weibull model. Setting the starting point of time recording also had an impact on the form of the hazard function, as shown in Figure 4.4. An additional benefit of shifting the start of recording from the onset to a later stage of laying is that hens with exceptionally low production, due to early culling, have not been included in the analysis. This may constitute a more straightforward method to exclude outliers than those considered in Chapter 3.

The interpretation of the regression coefficients corresponding to the explanatory variables in the parametric model demands careful consideration. Although positive values for the regression coefficients reveal a negative effect on the survival of the subject, within the context of this analysis, a positive coefficient denotes that a factor has an accelerating effect on the rate of lay and increases the intensity of lay. Hence, in this analysis, a positive coefficient is associated with meeting the breeding objective, which is an increase in egg production. In contrast, the progeny of sires, which are associated with negative regression coefficients, are expected to require additional time until reaching the production target.

Due to the apparent strong association between these two traits, it may be suggested that the development of survival models may have the potential to overcome some of the problems of the traditional selection for improving egg production. There is a series of arguments to advocate that this should be investigated. First, the normality assumptions that impair the estimation of genetic parameters for total egg number are not an issue under a survival model. Second, the outlier exclusion is handled by an explicit censoring rule when appropriate starting and terminating points are selected. Third, it appears that there is a considerable amount of genetic variation available for time traits. In this study it was shown that the heritability for the survival trait was comparable to that of the (un)transformed total number of eggs (survival: 0.16±0.08 - untransformed:. 0.11±0.04 - transformed: 0.22±0.04). The survival estimate may also be biased downwards, as it was derived from a sire model, while the other two were obtained from a multivariate model which accounted better for the ongoing selection. Fourth, time-to-event models allow the incorporation of both the persistency of lay and total egg production. In conclusion, survival models could constitute a versatile genetic tool to allow to identify good laying hens, not only based solely on their egg yield, but also considering their laying pattern. Moreover, as shown here, it is possible to construct events that give the desired distributional properties and so, a Weibull model can be used, which simplifies the estimation procedure and renders feasible large-scale analyses.

Despite the potential advantages of survival models, their development is still an ongoing process, which prevents them from being seamlessly incorporated into current breeding programs. One of the still unresolved issues is that the random model that is predominately used in genetic applications is not consistent with the animal model because it does not account for unobserved heterogeneity of environmental origin, since a residual effect is not directly included (Ducrocq and Casella, 1996). Although various approximate methods were used to estimate the individual heterogeneity, it was shown by simulation studies that parameter estimates may be biased (Damgaard and Korsgaard, 2006). Previous studies have suggested that although this bias impairs the prediction of selection gains, the ranking of sires appears not to be affected (Yazdi *et al.*, 2002). Furthermore, the estimation of best linear unbiased predictor for the random genetic effects is complicated, which makes these models less attractive from a practical point of view (Korsgaard *et al.*, 1998), unless a Bayesian/MCMC approach is adopted (Saebø and Frigessi, 2004)

Several of the issues noted above constitute fields of active research. Damgaard and Korsgaard (2006) developed a survival model that included a residual term in order to reduce the bias and improve the accuracy of prediction of breeding values. Another long-known constraint of using survival analysis is that a bivariate analysis between a time-to-even trait and a linear Gausian trait was not possible. Nevertheless, it has recently been indicated that this is feasible under a Bayesian framework (Damgaard and Korsgaard, 2006). However, these advances have not been yet diffused in a ready-to-use software for routine genetic evaluation.

Survival analysis provided a robust framework to define and investigate laying patterns, whose study was not straightforward with more conventional methods. This approach allows animal breeders address the longitudinal aspects of egg production and how these can be exploited to design more efficient breeding strategies. The development of frailty models to accommodate genetic effects is analogous to the introduction of a random effect in a linear regression framework as a way to account for correlation between clusters (Kleinbaum and Klein, 2002), which in a breeding context represents sire families. This observation puts the analysis presented in this chapter into the context with other methodologies to study longitudinal traits, such as random regression. The application of random regression models constitutes the topic of the following three chapters.

Chapter 5

The Application of Random Regression Models to the Genetic Analysis of the Monthly Egg Production in Turkeys and a Comparison with Alternative Longitudinal Models

5.1 Introduction

The analysis described in the previous chapter drew attention to the time-dependent features of the egg production. It was shown that the selection for body weight affects the intensity of lay, which in turn relates to the dynamics of lay. Nevertheless, there is limited information on the longitudinal aspects of the egg production, so before examining the genetics of egg laying dynamics *per se*, an investigation of the most appropriate statistical methodology was performed. This allowed me to build a framework to explore the genetics of egg production over time and perform an initial genetic analysis.

Random Regression (RR) models, have become a popular methodology for the genetic study of longitudinal data since the last decade. Literature is abundant with studies investigating the application of RR models in a wide variety of time-dependent traits (Schaeffer, 2004). However, the first and still one of the principal application field is in dairy cattle breeding, where RR models have underpinned the development of test day models for genetic evaluation.

Milk and egg production are analogous traits in that both change over time in a broadly comparable way and thus, when modelling the average production as a function of time, curves with similar shapes are generated. The success in the application of RRMs in dairy cattle attracted the attention of poultry breeders. Anang *et al.* (2002) reported that a RR model could be the most favourable model for analysing egg production data when compared to other longitudinal models, including a multi-trait one. Other studies investigated the efficiency of RR models for the genetic evaluation of egg related traits, such as fertility and hatchability (Sapp *et al.*, 2004).

The benefits from the application of RR models stimulate the investigation and the extension of their use in the genetic study of egg production. The source and the scale of variation in egg laying are not constant during the whole laying period. Changes over time can be expressed as a common component depending on time for all animals in the population that may be described with a fixed regression, and as an individual variation over time. Random regression is a method of modelling covariance functions that can be easily implemented into the genetic evaluations. So, RR models can detect both genetic and environmental factors that affect individual performance. Thus, a more accurate modelling of variance-covariance structure can lead to a more accurate prediction of breeding values (Huisman *et al.*, 2002). Furthermore, the function of genetic variance over time can be used to identify optimum points for selection and maximise its impact (Sapp *et al.*, 2004) or even alter the pattern of the genetic response (Schaeffer, 2004).

Therefore, the objective of this chapter was to investigate the application of RR models in the genetic analysis of egg production of turkeys. Furthermore, the same dataset was analysed using a multi-trait (MTM) and a repeatability (REP) model and results were compared. In order to assess the model efficiency of predicting missing values, a reduced dataset was used and the predicted values of the deleted records were compared for the three alternative models.

5.2 Materials and methods

The description of the population and the dataset can be found in Chapter 2. The production period was arbitrarily divided into five periods, each one including four weeks. In this way, each period corresponded approximately to the monthly egg production. So, the records used in the analysis were the sums of the eggs laid during a period of four weeks. All the analyses were performed using the AI-REML algorithm with the package DMU (Jensen and Madsen, 2006) and in all cases an animal model was considered.

5.2.1 Fixed regression

The data were analysed using repeatability model (REP), multi-trait model (MTM) and random regression model (RRM). The REP and RR models included a fixed regression part to account for the phenotypic trajectory of the average observations of the monthly egg production in different periods.

The shape of the laying pattern was modelled using the family of curves described by Ali and Schaeffer (1987). Although this model was initially introduced to describe the milking curve, its use can be extended to egg production data. Thus, the average egg production (y) of the population at time t is described by the following formula:

$$y = b_0 + b_1 \times \frac{t}{5} + b_2 \times (\frac{t}{5})^2 + b_3 \times log[\frac{5}{t}] + b_4 \times log[\frac{5}{t}]^2 \tag{5.1}$$

where b_0, b_1, b_2, b_3 and b_4 were the regression coefficients and t corresponded to each one of the five time periods (t = 1 - 5).

A combined fixed effect was fitted in all models to account for the factors year, hatch and pen. Since some classes of these parameters were missing, a superfactor was used to avoid having missing levels and accounting for the interactions between the factors.

5.2.2 Repeatability Model

The REP model treated the five-period summary measurements as repeated records. The model included a permanent environmental and an additive genetic effect and a fixed regression to model the phenotypic trajectory. Hence, the model to describe the egg production of period t (y_{ijt}) was:

$$y_{ijt} = S_i + FR_t + c_j + \alpha_j + e_{ijt} \tag{5.2}$$

where S_i was the i^{th} combined fixed effect, FR_t was the fixed regression terms given by equation 5.1, c_j and α_j were the random effects of the permanent environment and additive genetic effect respectively for the animal j (with c $\sim$ N(0, σ_c^2I) and $\alpha \sim$ N(0, σ_α^2A) respectively) and e_{ijt} was the residual (e $\sim$ N(0, σ_e^2I). This model assumes a genetic correlation of unity and independence of residuals across all periods.

5.2.3 Random Regression Model

An extension of the REP model was to include time functions in the random part of the model. The choice of the random regression function included two candidates: the Ali-Schaeffer function following the suggestion by Anang *et al.* (2001) and the Legendre polynomials. The model with the Ali-Schaeffer equation did not converged due to its complexity, so, the Legendre polynomials were finally used. This family of orthogonal polynomials has many advantages to be used as random regressions (Pool *et al.*, 2000).

The Legendre polynomials of order m were denoted as $\phi_m(w)$, where t is the standardized time period w_t using the following formula (Schaeffer, 2004):

$$w_t = \frac{2(t_i - t_{min})}{(t_{max} - t_{min})} - 1 \tag{5.3}$$

where, t_i was the i^{th} period, t_{min} and t_{max} corresponded to the earliest and latest period respectively (t_{min} = 1 and t_{max} = 5). According to the random regression model used, the egg production y_{ij} of period t was:

$$y_{ijt} = S_i + FR_t + \sum_{m=0}^{k} c_{jm}\phi(w_t) + \sum_{m=0}^{k} \alpha_{jm}\phi(w_t) + e_{ijt} \tag{5.4}$$

where S_i was the i^{th} combined fixed effect, FR_t was the fixed regression for month t given by equation 5.1, the third term represented the permanent environmental effect, the forth term represented the additive genetic effect of the j^{th} bird, and e_{ijt} was the residual term. Terms c_{jm} and α_{jm} represented the random regression coefficients on the Legendre polynomial functions.

The degree of the orthogonal polynomials, k was tested in order to determine the most proper combination. The likelihood of each model was compared with a log-likelihood test using the appropriate degrees of freedom, determined by the difference between numbers of model parameters (for each effect the degrees of freedom were: $\frac{1}{2}(k+1)(k+2)$, where k corresponded to the order of polynomials). Detection of non-zero eigenvalues of the corresponding eigenfunction of the covariance matrix provided further evidence for the necessary polynomials order (Meyer and Hill, 1997). Based on the log-likelihood test, the third order polynomial was the best (RR3), but the corresponding eigenvalue to the cubic regression was close to zero. So, the analysis was repeated for the second order polynomials (RR2) in order to compare the results.

Let matrix **G** be a 5x5 matrix of the estimates of variance for each period (in the diagonal) and the covariance between different periods (off-diagonal elements), it can be calculated by the covariance function (Kirkpatrick *et al.*, 1990):

$$\mathbf{G} = \Phi^T \mathbf{V} \Phi \tag{5.5}$$

where, Φ is a 4x5 matrix of the time covariates and **V** is a 4x4 matrix containing the covariance components of the random regression coefficients for the additive genetic effect (matrices Φ and **V** are 3x5 and 3x3 respectively when using second order polynomial). Likewise, a covariance matrix was computed for the permanent environmental effect (**C**). Finally, the residual covariance matrix (**R**) was the 5x5 identity matrix multiplied by the homogenous residual variance component. The total phenotypic covariance matrix (**P**) was the sum of the additive genetic, permanent environmental and residual covariance matrices (**P** = **G** + **C** + **R**).

The heritability (h^2) for time i and the genetic correlation (ρ) between time points j and k were defined as:

$$h^2 = \frac{g_{i,i}}{p_{i,i}} \tag{5.6}$$

$$\rho = \frac{g_{j,k}}{\sqrt{g_{j,j} \times g_{k,k}}} \tag{5.7}$$

where $g_{i,i}$ and $p_{i,i}$ were the diagonal elements of matrices **G** and **P** corresponding to the genetic and phenotypic variance for period i and $g_{j,k}$ was the element of the **G** matrix corresponding to the genetic covariance between periods j and k.

The standard error of heritability was calculated by extending the methodology proposed by Fischer *et al.* (2004), adapted to accommodate the output of DMU package. The formula used to estimate the variance of the heritability estimate for the i^{th} period was based on a Taylor series expansion and it was given by the following equation:

$$var(h^2) = (h^2)^2 \times (\frac{vg_{i,i}}{g_{i,i}^2} + \frac{vp_{i,i}}{p_{i,i}^2} - 2 \times \frac{COV(g_{i,i}, p_{i,i})}{g_{i,i} \times p_{i,i}}) \tag{5.8}$$

where $g_{i,i}$ and $p_{i,i}$, and, $vg_{i,i}$ and $vp_{i,i}$ were the diagonal elements of matrices **G** and **P** and **VG** and **VP** respectively. Matrices **VG** and **VP** correspond to the variance of **G** and **P** and were derived as the direct product of the variance-covariance matrix with the vector that included all the unique elements of matrix **P** (as in 5.5) for both genetic and permanent effects as long as the residual.

5.2.4 Multi-trait Model

A multi-trait model was also used in order to contrast it with the regression models. So, the egg numbers of the five sub-periods were treated as different traits and analysed simultaneously. Hence, the egg production y_{ij} of period k was:

$$y_{kij} = S_{ki} + \alpha_{kj} + e_{kij} \tag{5.9}$$

where S_{ki} was the i^{th} combined fixed effect, α_{kj} was the random additive genetic effect for the animal j ($\alpha \sim$ N(0,$\sigma_{\alpha}^2\mathbf{A}$)) and e_{kij} was the residual term ($e \sim$N(0,$\sigma_e^2\mathbf{I}$)). The additive genetic effect was given as the direct product of matrices **G**, the 5x5 matrix that describes the genetic variance-covariance between the five periods and **A**, the relationship matrix among the animals.

5.2.5 Model Comparison

In order to compare the predictive ability of the various models the data were divided into two parts. The first part corresponded to 80% of the total data and it was used to estimate parameters for the different models. This dataset was created as follows: within each generation, the first period was deleted for the first bird and so on. This was repeated for each group of five birds as they appeared in the dataset after sorting on generation, hatches and pen. Therefore, 2,400 observations were deleted, balanced in relation to the fixed effects. This corresponded to the 20% of the data and it was used to validate the model by comparing the predictions to the observed data using the mean square error of the difference. The rationale for this strategy was to study the model's ability to predict missing values, a situation often encountered in practice.

However, the MT model did not include a permanent environment as a second random effect, in contrast with the RR and REP models. This component is contained in the residual for each period and should be accounted for in the comparison. Due to the data separation protocol, residuals for the first, the second and so on periods were missing. The missing residual, conditional on the observed residuals of the other periods for an individual can be estimated via a multiple regression. So, the adjusted prediction was the sum of the fixed and the additive genetic effect plus the missing residual, estimated by the regression.

A further model comparison was performed only between RR2 and RR3, by assessing their prediction ability when only the last period was deleted. The objective of this second comparison was to detect if a model overfits the data. It was not possible to include MT model in this comparison, since it would not be possible to predict the deleted record from the four remaining ones.

5.3 Results

Phenotypic variances for egg production estimated from the three models are summarized in Table 5.1. For the REP model the estimates were assumed to be constant across the periods, while for the MT and RR, estimates were available for each of the five time points. The observed trend was that phenotypic variance increased along with the time periods.

Table 5.1: Phenotypic variance for all models and for all periods (RR2 = random regression model with second order regression; RR3 = random regression model with third order regression; REP = repeatability model and MT = multi-trait model).

Periods	RR2	RR3	REP	MTM
1	19.14	19.48	23.64	18.91
2	21.32	22.68	23.64	22.22
3	26.29	25.56	23.64	26.02
4	28.67	28.05	23.64	28.31
5	30.37	29.70	23.64	30.03

Heritability estimates are presented in Table 5.2 for all the three models. The RR3 had the higher estimates, with the exception of the fourth period. The heritability estimates from RR2 and MTM were very close for all the periods. The REP model had the lowest estimates from all models. In Table 5.3 the estimates of the ratio of the permanent environmental variance to the total variance (c^2) are also presented. The c^2 estimated from the RRM were close to each other regardless of the polynomial degree used, while the REP estimates were lower.

Table 5.4 presents the genetic correlations between egg numbers of different periods, estimated based on MT and on RR2 models. Although the values differed to some extent between the 2 models, the same pattern was observed. First, a weak correlation between the first and the middle stages of the production was observed. Second, the correlation between consecutive periods was strong. Third, the first and the last period were positively correlated. Phenotypic covariances are presented in Table 5.5. The phenotypic correlations were positive between all periods

Based on the heritability estimates from the MTM and RR models, a heritability profile was plotted by joining the estimates for each time point and thus covering the whole production period (Figure 5.1). This plot allows the visualization of the dynamics of the genetic variance over time.

Table 5.2: Heritability for all models and for all periods (RR2 = random regression model with second order regression; RR3 = random regression model with third order regression; REP = repeatability model and MTM: multi-trait model).

	heritability			
Periods	**RR2**	**RR3**	**REP**	**MTM**
1	0.12±0.03	0.13±0.04	0.04±0.01	0.14±0.04
2	0.05±0.02	0.11±0.03	0.04±0.01	0.06±0.03
3	0.07±0.03	0.08±0.03	0.04±0.01	0.07±0.03
4	0.07±0.02	0.05±0.02	0.04±0.01	0.05±0.03
5	0.07±0.03	0.08±0.03	0.04±0.01	0.10±0.03
TOTAL	0.07	0.08	0.04	0.07

Table 5.3: Estimates of the proportion of phenotypic variance explained by the permanent environment (c^2) for all models and for all periods (RR2 = random regression model with second order regression; RR3 = random regression model with third order regression; REP = repeatability model).

	c^2		
Periods	**RR2**	**RR3**	**REP**
1	0.45±0.03	0.50±0.04	0.45±0.03
2	0.56±0.03	0.58±0.03	0.45±0.03
3	0.60±0.03	0.64±0.03	0.45±0.03
4	0.64±0.02	0.70±0.02	0.45±0.03
5	0.65±0.03	0.69±0.03	0.45±0.03

The predictive ability was evaluated by comparing the difference between predicted and true values for the missing data. When data were deleted in a balanced fashion over periods, the RR3 model had the lowest MS error (MSE; MSE = 5.72), followed by the RR2 model (MSE = 13.84). The MTM model, though adjusted for the random effect

Table 5.4: Genetic correlation coefficients between all periods based on RR2 (upper triangle) and on MTM (lower triangle) (RR2 = random regression model with second order regression and MTM = multi-trait model) .

	RR2				
MT	Period 1	Period 2	Period3	Period 4	Period 5
Period 1	-	0.39	-0.09	-0.08	0.32
Period 2	0.17	-	0.87	0.79	0.47
Period 3	0.11	0.54	-	0.95	0.48
Period 4	0.05	0.62	0.77	-	0.71
Period 5	0.17	0.61	0.66	0.88	-

Table 5.5: Phenotypic correlation coefficients between all periods based on RR2 (upper triangle) and on MTM (lower triangle) (RR2 = random regression model with second order regression and MTM = multi-trait model) .

	RR2				
MT	Period 1	Period 2	Period3	Period 4	Period 5
Period 1	-	0.47	0.37	0.31	0.27
Period 2	0.48	-	0.61	0.57	0.47
Period 3	0.39	0.64	-	0.67	0.58
Period 4	0.34	0.54	0.68	-	0.66
Period 5	0.30	0.46	0.62	0.71	-

of the permanent environment, had a relatively large value (MSE = 16.31). The model with the largest MSE was REP (MSE = 18.89). Therefore, the comparison of the MSE suggested that the RR3 model was the most efficient to predict missing values. However, when the RR2 and RR3 models were compared using the reduced data set in which the last period was deleted for all hens, the results were different. In this case the RR2 model had the lowest MSE(MSE = 40.73), whereas the RR3 model appeared to have a larger

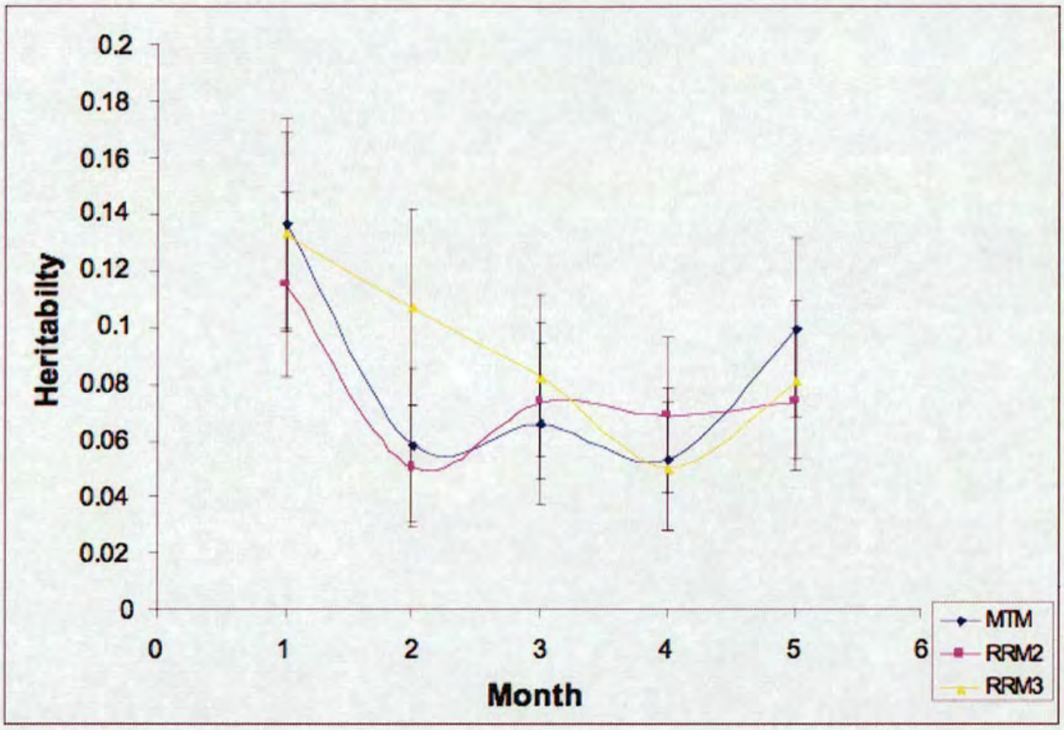


Figure 5.1: Heritability profile for all models (MTM multi-trait model; RR2 Random regression model using 2^{nd} order Legendre polynomials; RR3 Same as RR2 but using 3^{rd} order polynomials).

error (MSE = 67.43),

5.4 Discussion

The estimates of genetic parameters from the RR2 and RR3 model were comparable with the MT model, and both gave a detailed description of the dynamics of genetic variance in contrast with the REP model, which assumed a constant heritability and genetic correlation between periods. The present study showed that the heritability of 28 days egg production was high in the beginning of the laying period, decreased in the second period, remained constant for the rest of the time points, and increased again in the last stage of laying. The model comparison showed that both RR models, particularly the RR3 model, were more efficient in predicting missing values than the MTM or REP models but that the RR2 model was more robust to predict missing periods.

The use of the Ali-Schaeffer equation as the function for the fixed regression provided a robust tool to describe the trajectory of the average egg production (Anang *et al.*, 2001). In the current data set, the fit was perfect, because a 5-term equation was used to model 5 time points. However, a very good fit was also obtained when the same equation with the 5 time points was tested to fit for 10 or more time points (this is illustrated in the next Chapter, where 140 time points are fitted). This result provides evidence that the Ali-Schaeffer equation can be used satisfactorily to model the average egg production, with the benefit of being simpler than other models proposed, such as the Grossman *et al.* (2000) persistency model.

The REP model used in the current analysis was similar to the test-day models introduced by Ptak and Schaeffer (1993); however, it was unsatisfactory for the present data set, because more detailed analyses showed that the major assumption of the REP model did not hold. Results from the RR and MT models illustrated that the genetic correlations between egg productions of different periods varied from 0.1 to 0.9, whereas the REP model assumes a value of 1. Another assumption of the REP model is that genetic variance remains constant between periods, and this was not supported by the estimates derived from the RR and MTM models. In brief, the REP model offers a quick and simple approach for the genetic analysis of longitudinal data and has been used for the genetic evaluation of the egg production in poultry (Anang *et al.*, 2001), but the limitations stemming from the assumptions of the model makes it less preferable than other options.

The RR models offer an improvement over the REP model, because they allow the modeling of the genetic covariance between periods and are a development of covariance functions described by Kirkpatrick *et al.* (1994). Anang *et al.* (2002) concluded that RR was the preferred model for the genetic evaluation of egg production of laying chickens, and the current data extend this observation to turkeys when compared against the MTM model, which represents a more traditional approach to modeling repeated records over time. The comparison of genetic parameter estimates from the RR2 and MTM models showed that both models were equally effective to de- scribe the dynamics of the genetic variance over time. The general shape of the heritability profile obtained from the 3 models agreed with results from Anang *et al.* (2000, 2002). Similar trends are also observed for heritability of milk production using test-day models in dairy cattle

The RR models can deal with a large number of production periods with few parameters. In the present study, the total number of covariances for the MTM was 30, compared to 13 and 21 for RR2 and RR3 respectively. The model comparison using the first cross-validation strategy showed that both the RR had a lower mean square error than MTM. The lowest mean square error was obtained with RR3.

The superior prediction ability of the RR3 over RR2 could be associated with a larger number of explanation variables. Therefore a second cross-validation strategy was used to discriminate between RR2 and RR3. The second strategy was used to detect the ability of the two RR models to predict the egg production beyond the observed period, rather than predicting an internal missing value. Using the second strategy, RR2 gave the best fit, suggesting that the advantage of RR3 in the initial model comparison was a consequence of the larger number of explanation variables associated with RR3.

Further indications for rejecting the RR3 were provided by the different heritability profile when compared to RR2 and MTM and the larger standard errors. The latter suggests that the RR3 overfits the data, while further evidence of over-fitting was implied by the eigenvalue of the third order regression coefficient being close to zero, despite the fact that the RR3 had a lower log-likelihood value. Olori *et al.* (1999) used similar arguments when considering the appropriate order of RR when modelling lactation curves. Therefore, a trade-off seems to exist between the number of parameters and the model efficiency and so, determining the polynomial order requires consideration. It was concluded that RR2 was the most appropriate model in this dataset.

Apart from the appropriate polynomial order, the number of the time points that a RRM will fit is also crucial. Initially, a ten-period model was considered, but it failed to converge. Possibly, the underlying biological mechanism, involving overlapping ovipositions, interfered with the separation in ten periods. The egg number distribution within each period was more erratic and the approximation via the normal distribution less satisfactory.

One approach not followed in the current study was the use of transformations to reduce deviations from normality, even though this was considered in Chapter 3 for the whole laying period. The justification for excluding this approach was that simple data screening showed that a separate transformation would be required for each period. This would result in cumbersome evaluation procedures and would obscure inferences.

In conclusion, the application of RR in egg production of turkeys appears to be promising. It can effectively model the laying procedure even when missing values exist, as highlighted in the current study. The implementation of RR allows the genetic evaluation of egg production on a monthly basis and could provide helpful information for breeders to optimize selection strategies. Nevertheless, in view of the rapidly changing heritability over the initial period, use of more time points may be warranted, in order to increase the accuracy of estimates and thus, to derive the full benefits of modelling the genetic variation over time for providing a more reliable framework for breeders. This analysis constitutes the topic of the following Chapter.

Chapter 6

Genetic analysis of daily egg production in turkeys under a Bayesian framework

6.1 Introduction

In the previous chapter the advantage that random regression (RR) models offer over repeatability and multiple trait models was shown. The division of the laying period into intervals allowed to investigate the longitudinal aspect of egg production. Hence, it can be suggested that increasing the number of sub-periods would improve the accuracy of the modelling, since even minute changes will be captured. However, a limitation in the number of sub-periods seems to exist in the analysis of egg production data in turkeys using RR linear Gaussian model. In the analysis described in the previous chapter it was not possible to obtain convergence when periods shorter than a month were fitted in the RR model.

Under an alternative modelling approach, time periods can be set to cover only a single day. Thus, egg laying is converted to a binary trait, where 0 corresponds to none egg laid and 1 to an egg laid. Therefore, daily egg production over time can be modelled using a liability RR model. Sorensen *et al.* (1995) introduced a threshold model for categorical responses using a Bayesian approach, which was further developed by Sorensen and Gianola (2002). Rekaya *et al.* (1998) extended the model by including a random regression term for the study of longitudinal binary responses. Successful model implementations were found in studies in dairy cattle performed by Rekaya *et al.* (2003), Heringstad *et al.* (2003) and Chang *et al.* (2004) for the genetic analysis of clinical mastitis and by Averill *et al.* (2006) for fertility evaluation. Such conclusions are the stimulus for adaptation of this model for studying the structure of the genetic covariance of daily egg production in turkeys.

The objectives of this study were, first, to estimate genetic parameters for the daily egg production over the whole laying period in turkeys using a longitudinal threshold model under a Bayesian framework; and second, to explore the possibility of using the information obtained in order to investigate alternative selection strategies for the genetic improvement of laying turkeys.

6.2 Materials and methods

A detailed description of the dataset can be found in Chapter 2. The data used in this analysis consisted of the daily egg production of hens over 140 days.

6.2.1 Genetic model

Since daily egg production was treated as a binary trait, it was postulated that the observed record y_i of the i^{th} day in the production period is associated with an unobserved underlying variable, the liability λ_i ($\lambda_i \sim \mathrm{N}(\mu_i, \sigma^2)$). So, if $\lambda_i \geq T$ then an egg is laid in the i^{th} day ($y_i = 1$) and if $\lambda_i < T$, then no egg is laid ($y_i = 0$), where T is an unknown fixed threshold (Sorensen and Gianola, 2002).

The term μ_i represents the mean liability of the population for each production day (i=1...140) and in order to account for the population trajectory over time, the formula introduced by Ali and Schaeffer (1987) was fitted as a fixed regression. Although this function was introduced for describing the lactation curve, it is also adequate to model the pattern of egg laying (Anang *et al.*, 2001). Some experimentation with the current dataset suggested that a better fit was achieved when using Ali-Schaeffer for fixed regression compared to Legendre polynomials.

In order to account for individual liability as a deviation from average performance, random regression terms were fitted for both the additive genetic effect of the sire and a second random term peculiar to each record that included the remaining genetic and permanent environmental effects, which will be referred to as a permanent environmental effect. Legendre orthogonal polynomials were used as random regression functions because of their benefit of increasing the mixing properties of the Markov Chain Monte Carlo (MCMC) process (Chang *et al.*, 2004). The Legendre polynomials of order *m* were denoted as $\phi_m(wt)$, where *wt* was the standardised time period using the following formula (Schaeffer, 2004):

$$w_t = \frac{2(t_t - t_{min})}{(t_{max} - t_{min})} - 1 \tag{6.1}$$

and where, t_t was the t^{th} period, t_{min} and t_{max} corresponded to the earliest and latest period in the present study respectively (t_{min} = 1 and t_{max} = 140). The order of the polynomial m was set equal to two, based on the results from Chapter 5.

Thus, the unobserved liability λ_i of a hen j of sire k to lay an egg in day t was:

$$\lambda_{ijkt} = YHP_i + b_0 + b_1 z_t + b_2 z_t^2 + b_3 log[z_t^{-1}] + b_4 log[z_t^{-1}]^2 + \sum_{m=0}^{2} c_{jm}\phi(w_t) + \sum_{m=0}^{2} s_{jm}\phi(w_t) + e_{ijkt} \tag{6.2}$$

where YHP_i was the combined fixed effect for the factors year, hatch and pen, z_t = t/140 and b_0, b_1, b_2, b_3 and b_4 were the regression coefficients of the Ali-Schaeffer fixed regression for day t. Terms c_{jm} and s_{km} represented the remainder genetic plus the permanent environmental and the sire genetic random regression coefficients of hen j and sire k, respectively, for Legendre polynomials $\phi_m(w_t)$ of order m, where w_t was the standardised day t (Kirkpatrick et al., 1990). The residual effect, e_{ijkt}, was assumed to have constant variance over time equal to one ($e \sim$ N(0,1)). The sire model was preferred to a full animal model due to the computation benefits. Sires were considered to be related.

A Bayesian MCMC approach was used to fit the above model 6.2 to data. Uniform, independent prior distributions were assigned to fixed effects and to all fixed regression coefficients, while the random regression coefficients were assumed to follow a multivariate normal prior (Chang *et al.*, 2004). Both permanent and additive genetic variances were assumed to follow independent scaled inverse X^2 prior distribution.

Draws from the fully conditional posterior distributions were obtained using a Gibbs sampler with data augmentation (Sorensen *et al.*, 1995). A single long chain of length 1,050,000 with a burn-in period equal to 50,000 was used. A little experimentation indicated that such a chain led to small Monte Carlo variances of features of posterior distributions. Convergence was monitored visually by inspection of traceplots.

6.2.2 Inferences of genetic parameters

The covariance matrix **S** for the sire genetic effects was computed as:

$$\hat{\mathbf{S}} = \Phi \mathbf{V} \Phi^T \tag{6.3}$$

where $\hat{\mathbf{S}}$ was the estimate of **S** with order 140x140, Φ was a 140x3 matrix of time covariates and **V** was a 3x3 matrix containing the variance components of the intercept and the random regression coefficients for sire genetic effect. Likewise, a covariance matrix was computed for the permanent environmental effect and a residual covariance for the 140 laying days was constructed from a 140x140 identity matrix multiplied by the homogenous residual variance.

These matrices were used to infer the genetic parameters on the underlying scale. Thus, the total phenotypic covariance matrix (**P**) was the sum of the sire, permanent environmental and residual covariance matrices. The liability heritability (h^2_λ) for day i and genetic correlation (ρ_λ) between days j and k were defined as:

$$h^2_\lambda = \frac{4\mathbf{S_{i,i}}}{\mathbf{P_{i,i}}} \tag{6.4}$$

$$\rho_\lambda = \frac{\mathbf{S_{i,j}}}{\sqrt{\mathbf{S_{i,i}S_{j,j}}}} \tag{6.5}$$

Moreover, a new trait was defined: the cumulative liability ($C_{x,y}$) that corresponded to the sum of liabilities between days x and y ($1 \leq x < y \leq 140$). The heritability of $C_{x,y}$ between days x and y at the level of the liability was estimated as follows:

$$h^2_{\lambda C_{x,y}} = \frac{4c^t Sc}{c^t Pc} \tag{6.6}$$

where **c** was an indicator vector with 140 elements corresponding to the days of laying period, with $c_i = 1$, when $x \leq i \leq y$ and $c_i = 0$ otherwise. The heritability of the total liability (h^2_λ) corresponded to $x = 1$ and $y = 140$.

The genetic correlation $\rho_{\lambda(C_{x,y,T})}$ between the cumulative liability $C_{x,y}$ from day x to day y

and the total liability T, was defined as:

$$\rho_{\lambda(C_{x,y},T)} = \frac{c^t St}{\sqrt{(c^t Sc)(t^t St)}} \tag{6.7}$$

where **c** and **t** were indicator vectors with 140 elements corresponding to the laying days and **S** was the sire covariance matrix. The values of vector **c** elements were set as described in equation 6.6, while all vector **t** elements were equal to one, since they corresponded to the whole period.

The correlation between cumulative and total liability ρ_{CT} was used to assess the efficiency of indirect selection for cumulative liability of egg production from day x to day y in lieu of the direct selection for total liability. Following Falconer and McKay (1996) indirect selection is more efficient than direct selection when the following inequality is valid, given that the intensity of selection remains the same in both cases:

$$h_i \times \rho_{id} > h_d \tag{6.8}$$

where, h_i and h_d are the square root of the heritability of the indirectly and directly selected trait respectively estimated from 6.6, and ρ_{id} is the genetic correlation between the two correlated characters, estimated from 6.7.

All combinations of starting (x) and finishing (y) days were examined to identify the intervals where indirect selection is more efficient than direct selection, on both the liability and observed scales. Since the problem space consists of 140x140 states, a two-step procedure was followed. First, an initial screening of problem space was performed by substituting the function terms in 6.8 with the estimates obtained from posterior means. This indicated the candidate periods for which indirect selection appeared to be more effective than the direct on average. The second step was to calculate again 6.8 for each of the available MCMC iterations and compute the ratio of those satisfying the inequality versus the total number of the iterations. This probability indicated the effectiveness of indirect selection using the partial production of a specific period compared to the direct selection for total liability.

Since a liability model was used, all the variance component estimates referred to the underlying scale and hence results had to be translated to the observed scale. Robertson and Lerner (1949) showed that in the case of a character having only one threshold the heritability on the observed scale (h^2_{obs}) is equal to:

$$h^2_{obs} = \frac{[f(x_p)]^2}{p \times (1-p)} \times h^2_{\lambda} \tag{6.9}$$

where the h^2_{λ} is the heritability on the liability scale, p is the proportion of hens laying an egg each day with the product $p(1-p)$ corresponding to the phenotypic variance in the observed scale ($\sigma^(2)_{obs}$) and $f(x_p)$ is the height of the standardised normal curve at the position of the threshold x_p (Lynch and Walsh (1998) and Mrode (2005)). Equation 6.9 can be generalised to convert the covariance matrices on the observed scale, if the heritability in the liability scale is written as the ratio of the genetic and phenotypic variance on the liability scale. Thus, the covariance matrix for the sire on the observed scale ($\mathbf{S_{obs}}$) was obtained as follows:

$$\mathbf{S_{obs}} = \mathbf{F_z} \mathbf{S}_{\lambda} \mathbf{F_z} \tag{6.10}$$

where $\mathbf{S}_{\lambda}$ is the covariance matrix for the sire effect on the liability scale, and $\mathbf{F_z}$ a 140x140 diagonal matrix with i^{th} diagonal element corresponding to the i^{th} day of the period equal to:

$$\mathbf{F_z}(i,i) = \frac{1}{\sigma_{\lambda i}} \times \frac{f(x_{pi})}{\sqrt{p_i \times (1-p)}} \tag{6.11}$$

where $\sigma_{\lambda i}$ is the phenotypic variance, p_i the proportion of laying hens and $f(x_{pi})$ is the height of the standardised normal curve for the i^{th} day. Equation 6.10 was also applied to obtain the covariance matrix of the permanent environmental effect on the observed scale.

Finally, the phenotypic covariance matrix on the observed scale was also obtained as follows: the diagonal elements were set to be equal to $p(1-p)$ and the off-diagonals were set to be equal to the sum of the corresponding elements of the sire and permanent environmental matrices. Using this scaling method, the modified covariance matrices were used in equations 6.4, 6.5, 6.6, 6.7 and 6.8 in order to estimate the heritability and genetic correlation, and to investigate direct and correlated response to selection on the observed scale. As follows from 6.5 and 6.10, the genetic correlation was the same on both scales.

The availability of all MCMC iterations made feasible the estimation of moments of the posterior distribution for each estimated parameter, including means, variances as well as coefficients of skewness and kurtosis. All parameters were estimated as the mean of their respective posterior distribution.

6.3 Results

6.3.1 Fixed regression

The fixed regression coefficients, after being transformed to the observed scale, were used to model the average daily egg production. Hence, the predictions were plotted as a function of time and compared against the raw data for hen-day production (6.1). The dependent variable was expressed as the hen-day production, i.e. the probability of a hen laying an egg on a specific day of the production period. In the beginning of the laying period the fitted values from the model agreed well with the raw means, but in the second half the model appeared to underestimate the average daily production. This deviation is addressed in the discussion section of this Chapter.

6.3.2 Genetic parameters for the daily egg production

The additive genetic variance for each day was estimated both on the liability and on the observed scale. The coefficients of skewness and kurtosis of the posterior distribution were calculated for each day using all the saved MCMC samples and were plotted against all laying days in Figure 6.2. This plot indicates that the posterior distribution of the heritability estimates for all time points is characterised as right-skewed and leptokurtic

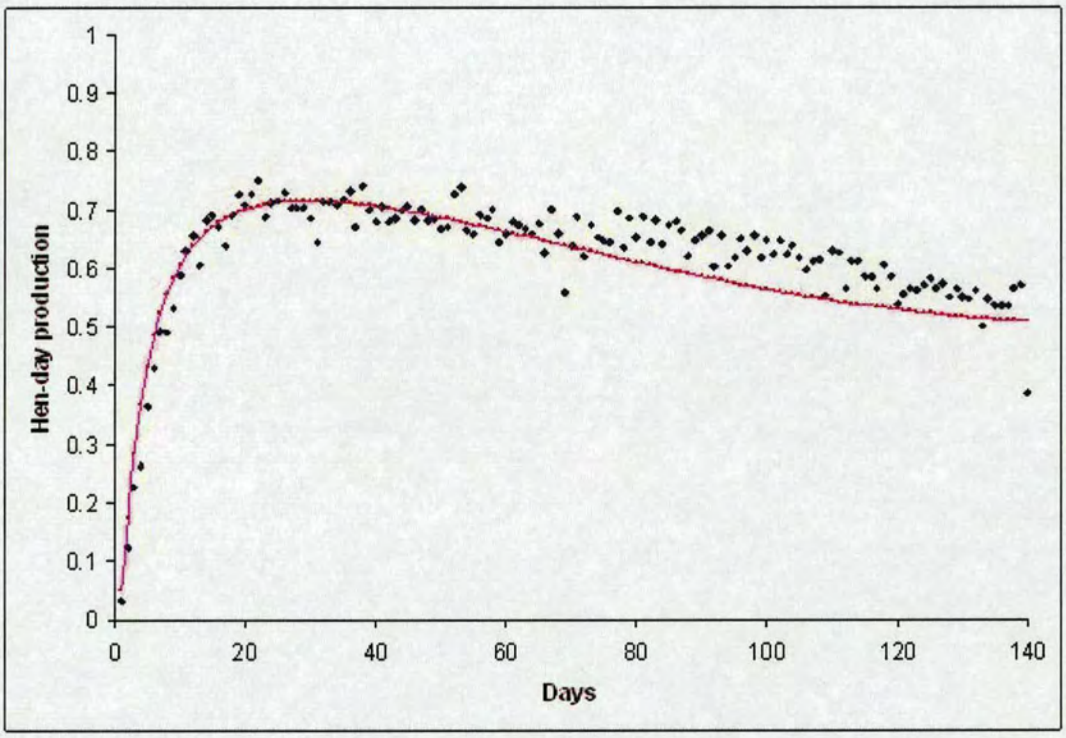


Figure 6.1: Fit of the fixed regression modelling the average daily laying curve (hen-day production) over the raw data (points).

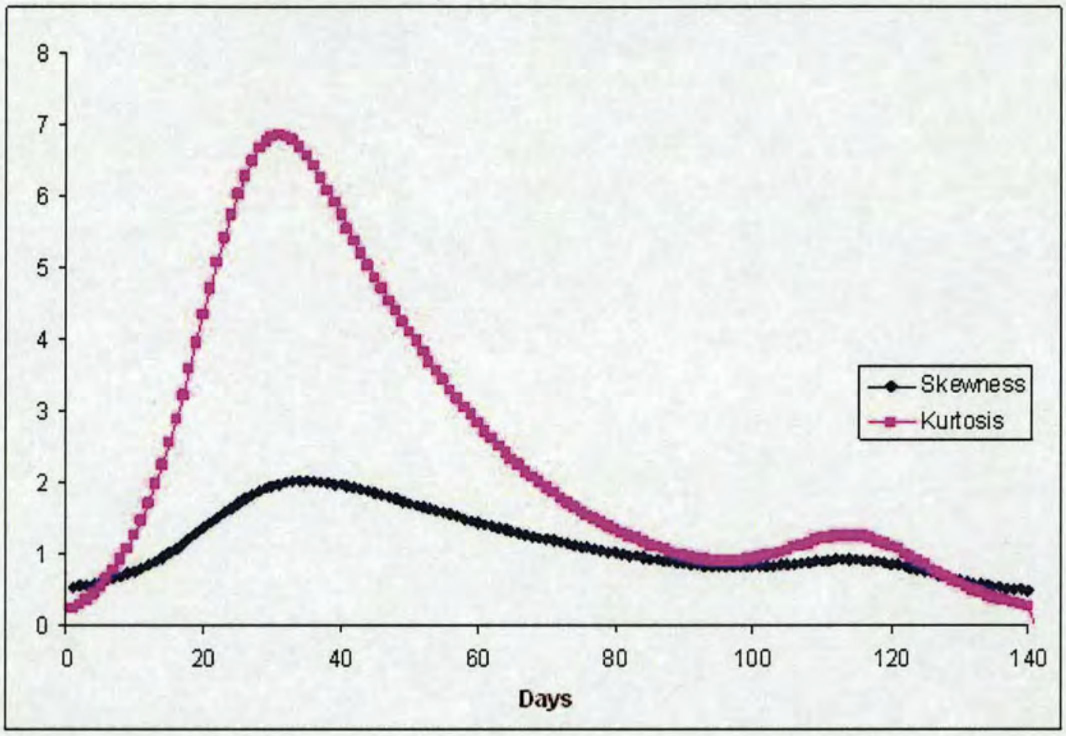


Figure 6.2: Skewness and kurtosis of the distribution of daily heritability on the observed scale (the curve shape was similar for both moments in the liability scale).

on both the liability and the observed scale. Figures 6.3 and 6.4 display the heritability plot during whole production period on the liability and on the observed scale respectively. A common observation was that the heritability increased at the onset of laying but rapidly decreased, reaching a minimum level in the first month, and towards the end of the production period rose again.

The inconsistency between the heritability estimates on liability and on observed scale (Figures 6.3 and 6.4) during the first production days is explained by the low frequency of daily egg production at this period and how this affects the estimation of heritability on the observed scale (i.e. p in equation 6.9 was small). As at the onset of laying, the hen-day production was low, so the heritability on the observed scale was also low, despite the estimated heritability on the liability scale being at its maximum.

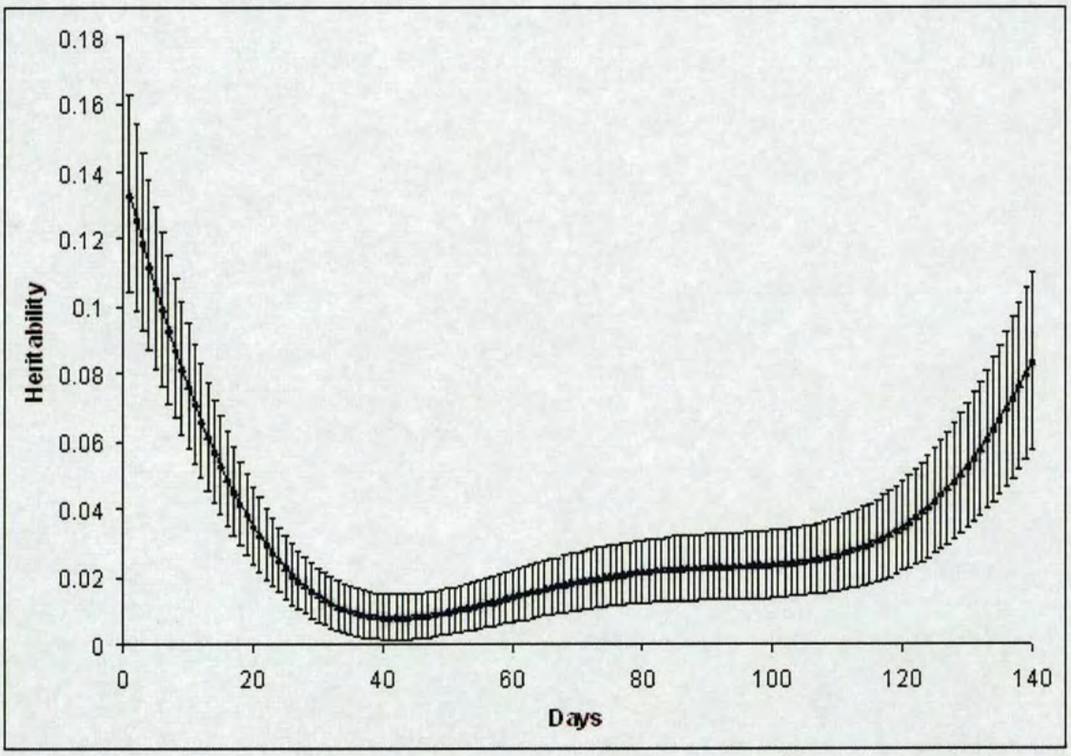


Figure 6.3: Plot of the daily heritability in the liability scale, including the standard deviation of the estimates.

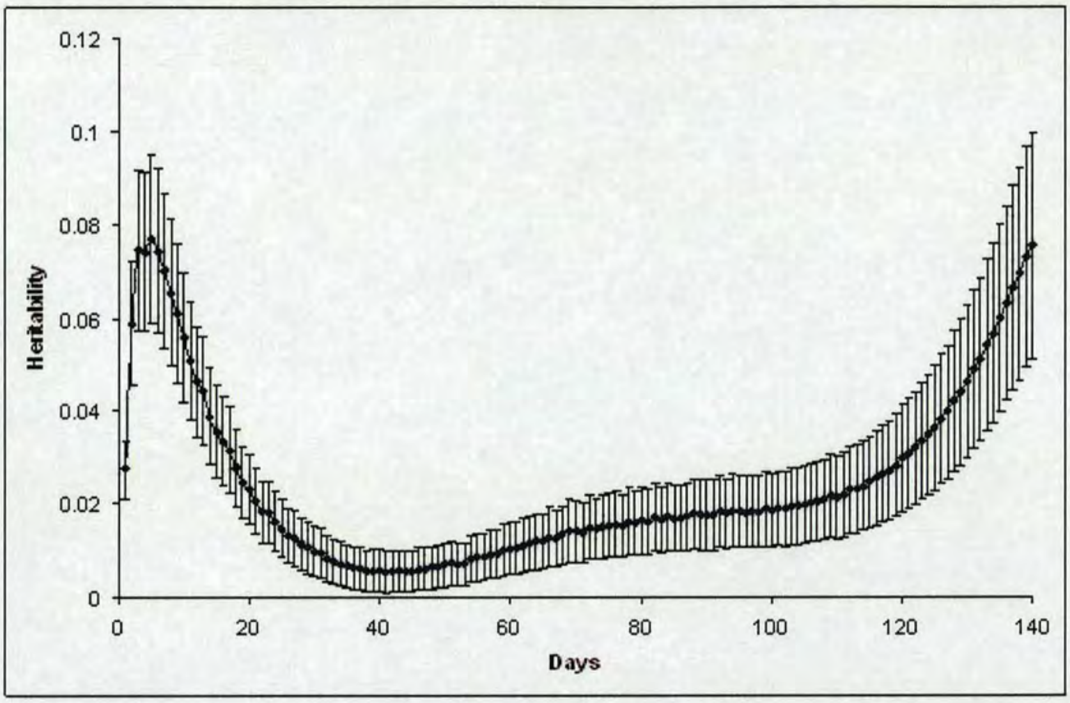


Figure 6.4: Plot of the daily heritability in the observed scale, including the standard deviation of the estimates.

A surface plot of the genetic correlation between all days of the laying period is presented in Figure 6.5. The estimates on the liability and observed scale were identical (see section 6.2.2). A high genetic correlation was observed for consecutive periods. The first days of laying were negatively correlated with days around the middle of the production period. In contrast, the first and last days of the laying period were positively correlated at the genetic level, providing evidence that the persistency of lay has a genetic basis. The same pattern was observed for the phenotypic correlation.

6.3.3 Cumulative liability/egg production from day one to day x

The posterior distribution of the heritability of cumulative egg production from day one to day *x* ($x \leq 140$) was unimodal, right-skewed and leptokurtic on both underlying and observed scales. Standardised moments of the posterior distribution were plotted over time

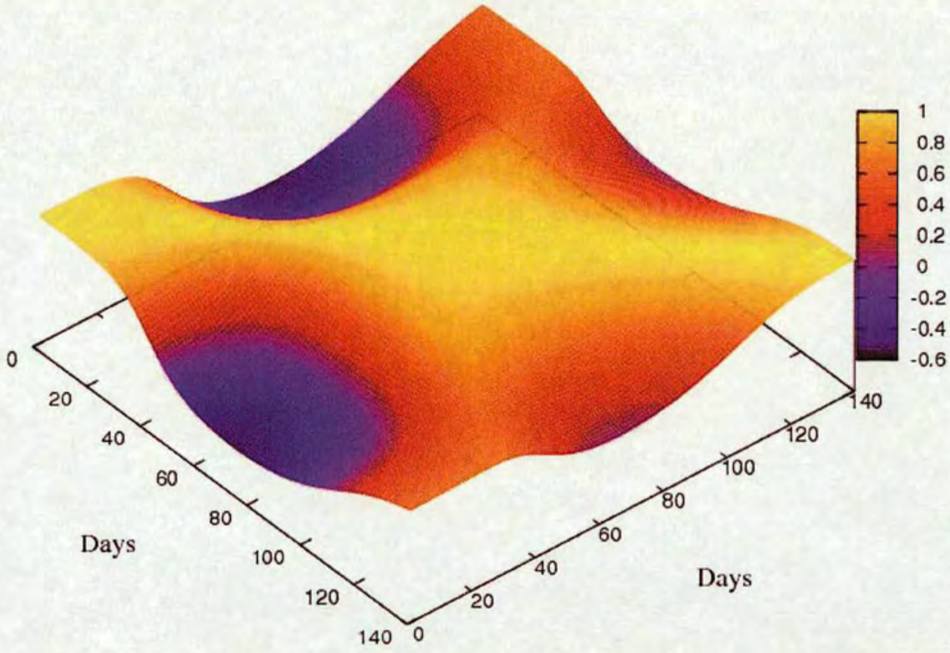


Figure 6.5: Surface plot of the genetic correlation between all days of laying period.

and it was concluded that the curve shape was similar to the plot presented in Figure 6.2 (results are not shown). Figure 6.6 presents the heritability plot of the cumulative liability starting from day one and extending up to the last day of the laying period. The peak of the heritability was observed for the cumulative liability that covered the first twenty laying days. The same pattern was observed on the observed scale (results not shown), but the absolute values of estimates were lower due to the scaling applied.

The heritability of total liability was found to be equal to 0.09 ± 0.04. This estimation, after being translated to the observed scale, corresponded to the heritability of the total egg number and it was estimated to be 0.06 ± 0.04.

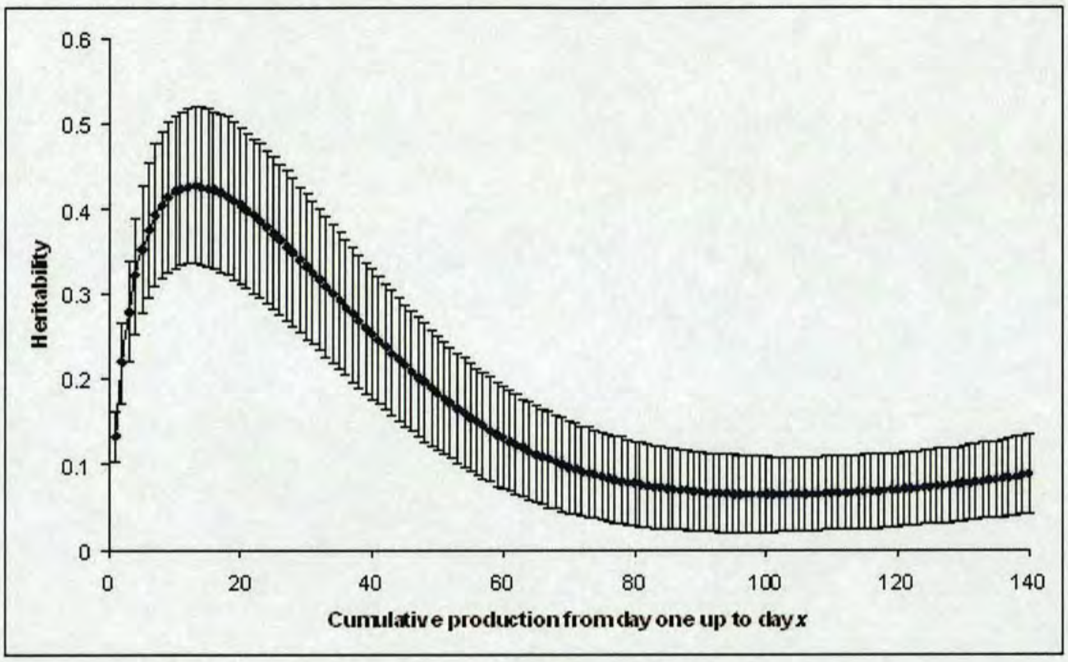


Figure 6.6: Plot of the heritability estimates, including the standard deviation, of the cumulative egg production from day one up to day x (x $\leq$ 140) in the liabiltiy scale.

6.3.4 Cumulative liability/egg production & indirect selection

The means of the posterior distributions for total liability and egg production were used as the criterion in equation 6.8 to identify efficient selection windows. Figure 6.7 presents the result of this search on the liability scale. The optimum time periods are visualised as light areas (with yellow colour) on the surface plot. The graph shows that indirect selection to improve total egg production appeared to be more efficient when selecting for cumulative egg production during the last days of the production period. The systematic search in the second half of the laying period (including all periods starting from day x and finishing to day y with $71 \leq x < y \leq 40$) suggested that the probability that an effective selection window existed during this time frame exceeded 0.85. The combination for which the indirect selection appeared to be the most effective is the one that corresponded to the partial production between days 120 and 140. This observation was validated for 72% of all available draws from the posterior distribution.

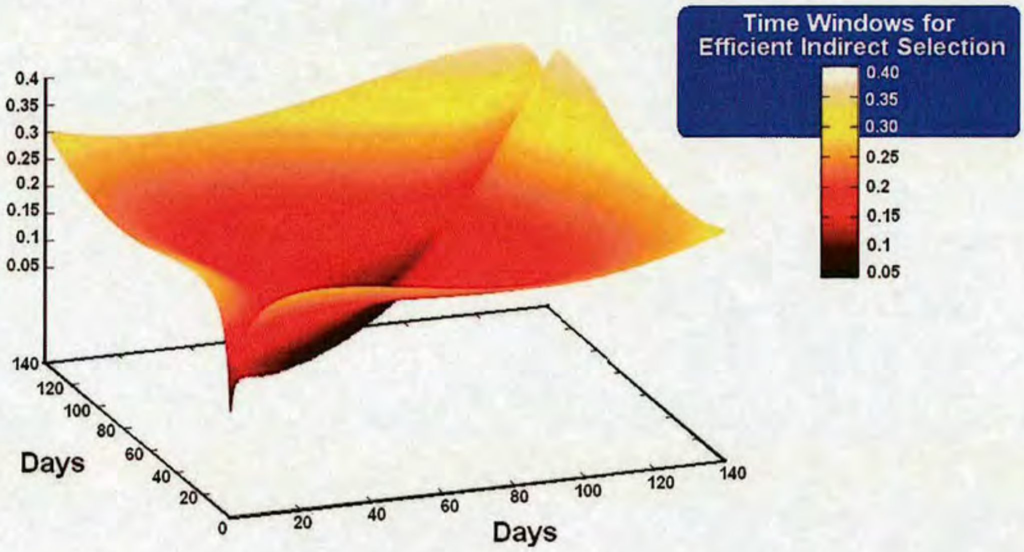


Figure 6.7: Search for optimum time windows (marked as light-coloured, yellow) for which the indirect selection using the cumulative production between two days, corresponding to x and y axes of the contour plot, is predicted to be more efficient than the direct selection for the total egg number in the liability scale.

Figure 6.7 indicates that a selection window may also exist for partial egg production whose recording starts from the first day. The probability of observing a selection window during the first half of the laying period (all combinations of periods starting from day x and finishing day y with $1 \leq x < y \leq 70$), in which the indirect selection was predicted to be more efficient than the direct selection, was 0.65 on the liability scale and 0.63 on the observed scale. In this part, the most efficient combination on the liability scale appeared to be the periods for which recording of eggs started at the first day and finished between days 14 and 34. In Figure 6.8 this interval corresponds to the x values (days) for which the curve surpasses the straight line corresponding to the square root of the heritability of total liability (right-hand side term in equation 6.8). The occurrence of this window was validated for 50% of all available draws from the posterior distribution, as Figure 6.9 illustrates. The combination for which the indirect selection appeared to be the most effective during the first half of the laying period was the cumulative liability

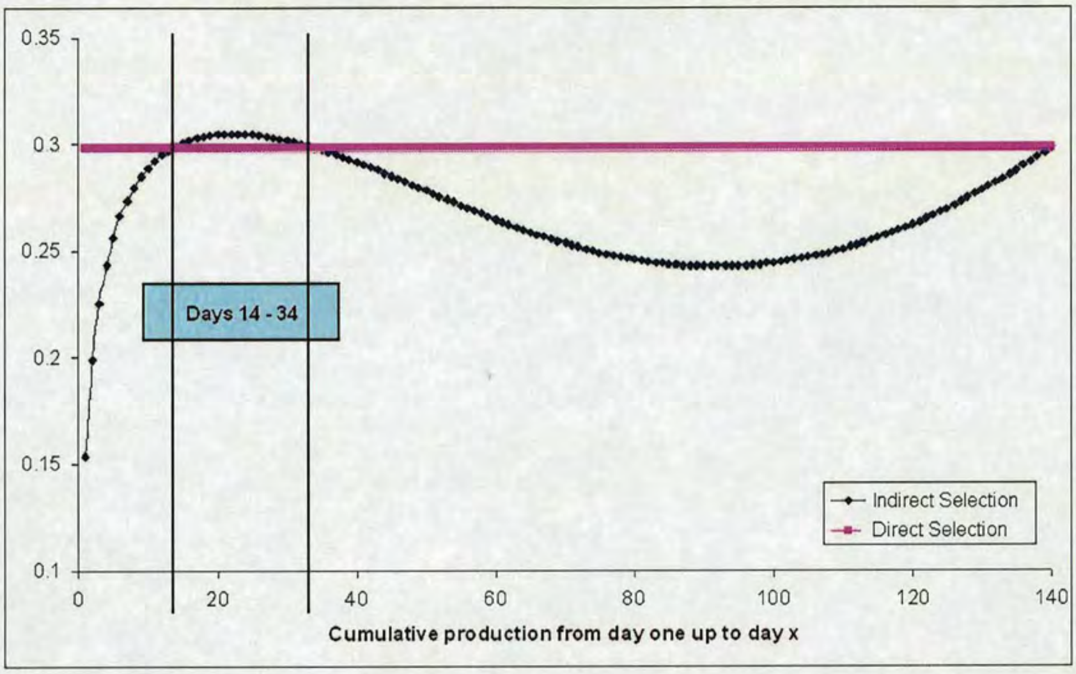


Figure 6.8: Comparison of the indirect selection for cumulative production beginning from day one up to day *x* and the direct selection for total egg number. Partial egg production starting from the first day and finishing between days 14 and 34 is the only interval when the indirect selection is predicted to be more efficient.

between day one and twenty-two. This was observed for 53% of all available draws from the posterior distribution.

6.4 Discussion

The fixed regression described satisfactorily the average daily egg production for the first part of the laying period. The random regression provided heritability and genetic correlation estimates on a daily basis that allowed a detailed insight to be obtained of the dynamics of the genetic parameters of egg production for the whole production period on the liability scale. A selection window, where the indirect selection using cumulative egg production as a criterion appeared to be more efficient than the direct selection for

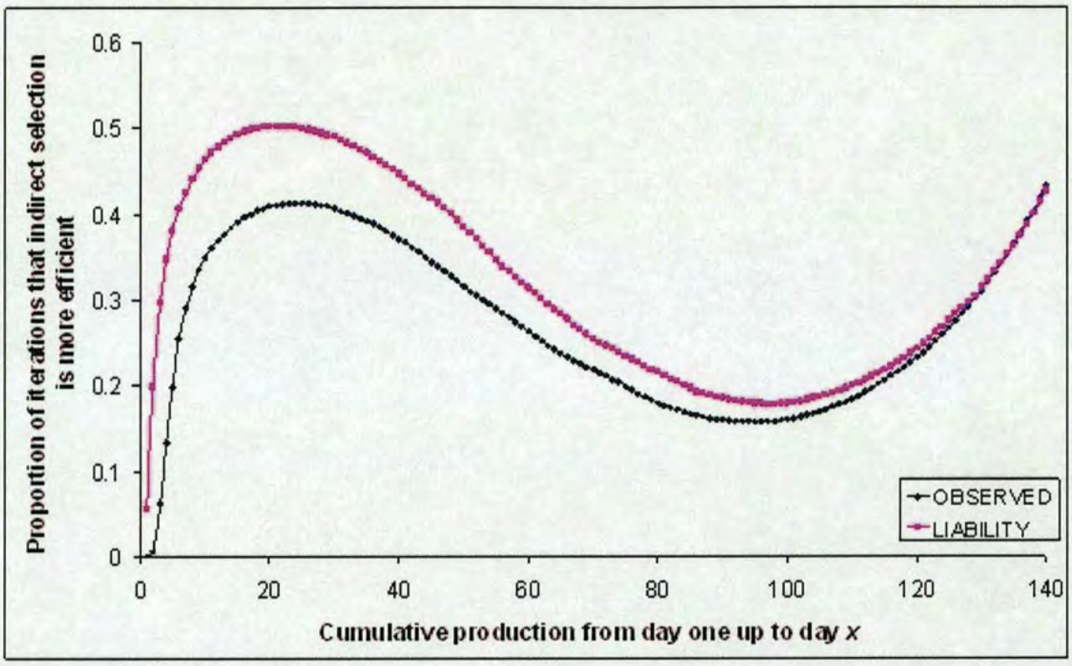


Figure 6.9: Likelihood of an efficient selection window existing when the cumulative production from day one up to day *x* ($x \leq 140$) is used for indirect selection for the total egg number in both the liability and observed scale.

total egg number, was observed for both the first and the last month of production, with the latter appearing to be more favourable due to the higher efficiency and likelihood of occurrence. However, this potential advantage needs to be weighted against other factors that govern the design of breeding programmes.

6.4.1 Liability and observed scale

In this Chapter a liability model was used, and therefore, results are applicable on the underlying basis and careful consideration is required when they are interpreted on the observed scale. Nevertheless, the adoption of a liability model may offer certain advantages for the genetic analysis of egg production.

Daily egg production is perceived as a binary character, despite being a quantitative trait presumably regulated by a large number of loci. The use of the liability scale allows accounting for these polygenic effects on a continuous scale and therefore, it may be suggested that this approach provides a model closer to the underlying biological mechanisms. Furthermore, because the liability of daily egg production is assumed to follow the normal distribution, the departure from normality that hampers the analysis on the observed scale (see Ibe and Hill (1988) and Chapter 3) is no longer an issue. Hence, the use of the unobserved scale may constitute a more biologically meaningful approach than the power transformations for removing the bias on the estimation of the genetic parameters, due to the discrete character of egg records on the observed scale.

This becomes evident in the case of a bivariate analysis of egg production and body weight. Because of the departure from normality, the common approach is to apply a power transformation, as in Chapter 3. So, the genetic correlation estimated is in fact that between the transformed egg production and body weight, which is assumed to reflect more closely the true relationship between the traits. However, if the liability of egg production is considered, then no transformation on data is required. Furthermore, it has been suggested that the underlying scale may provide a more accurate model of the biology of laying. So, it may be postulated that the genetic correlation between the liability of egg production and body weight is a more precise estimation of the true genetic relationship between the two traits.

Despite the advantages of the use of the liability scale, a major drawback is that all the estimates need to be translated back to the observed scale. For this purpose, equation 6.10 is used, but this provides only an approximation. Therefore, careful interpretation of results is recommended when using results from a liability model in order to make inferences for selection on the observed scale.

6.4.2 Laying patterns

The fixed regression model accounted for hens culled, corresponding to hen-day production. Despite the good fit for the first half of the laying period, toward the end of the production period the model appeared to underestimate raw means. Explanations could be either an inefficiency of Ali and Schaeffer equation to adjust to the shape of the laying curve or as if the raw means are biased upwards. To test the model capability, a simple multiple regression was applied to raw means and the fit of the model improved, suggesting that the function can effectively describe the shape of the laying curve. The alternative explanation involves biased raw means, due to genetic trends from selection for which the predicted egg production did not account. The number of birds removed from production was larger for the first half, and the culled birds were characterised by low yield. Despite this discrepancy, the fixed regression term appeared to model satisfactorily the average egg laying in turkeys, offering a robust alternative to previous approaches (for a review see Fairfull and Gowe (1990)),

The high genetic variance and heritability in the initial period most likely reflected the variance observed for the onset of laying and my decision not to correct data for the age at first egg, as it is illustrated in Figure 6.3. The period between twenty and sixty days was characterised by a reduced heritability, but this coincided with the stage when hens' laying rate is maximised (Figure 6.1). Furthermore, during this period the proportion of the permanent environmental effect was increased in contrast with the sire component that was reduced. So, the heritability drop was a result of the reduced contribution of genetic variance and the respective increase of the permanent environmental effect because the majority of hens were laying well. However, towards the end of lay, when the rate of lay was reduced, sire variance increased and a boost in heritability was observed.

The trends of the daily heritability were consistent with the study performed in the previous chapter using a REML random regression and a multi-trait model examining the monthly egg production. The current observations on laying patterns were also in compliance with the findings of Anang *et al.* (2002), in which they reported the application of RRMs in the monthly egg production of chicken. Fairfull and Gowe (1990) compiled a consensus table from earlier studies presenting the heritability of individual laying stages in chicken, indicating an agreement with our results about the patterns of lay. The patterns of genetic correlation were also consistent with the results presented in Chapter 5.

6.4.3 Indirect selection using partial production

The use of part records as a criterion to indirectly select for egg production has been investigated in the past for egg laying chickens (for a review see Ayyagari *et al.* (1980) and McMillan *et al.* (1986)). However, in all the previous studies each period was analysed separately and hence the level of accuracy may not has been optimal. A detailed model, like the one used in this study, may shed light on the complex dynamics of egg laying.

The search for efficient selection windows showed that intervals towards the end of the production period appeared to be more efficient. The high genetic correlation between the cumulative production of the last days of laying and the full production period could serve as an explanation (Figure 6.5). There, it is shown that the start and the end days of the laying period were positively correlated. It can also be implied that the persistency of lay provides an additional explanation, since birds that perform well in the last stage of lay are likely to have an analogous egg yield in the previous periods resulting in a high total egg output. The importance of persistency has also been suggested from the survival analysis in Chapter 4, where it was indicated that the production of the last laying stage was strongly correlated to the total yield.

Despite the advantage of using the cumulative production of the last production days, the efficiency of a breeding program is influenced by many factors, including the timing of selection. In this context, waiting for the last days of the production to perform the selection may not be the best practice, because it prolongs the generation interval. Hence, the benefit from the more efficient selection should be compared with the loss of progress due to longer periods between consecutive generations.

For this reason the case of the cumulative production starting from day one up to day x was also considered, although it appeared not to be the most efficient selection window. The periods, for which the probability of occurrence of an efficient selection windows exceeds 50%, are the partial records covering the period from day one to days between the 14^{th} and 34^{th} (Figure 6.9). This selection window may be beneficial for breeders, since our results implied that the selection based on the partial production of the first days may be at least as efficient as waiting to obtain the total egg number. This can accelerate the genetic progress, since it allows earlier selection and thus, minimising the generation interval. Additional benefits from earlier selection can be that the maintenance cost of the nucleus flock can be reduced, since the recording of the egg production on a daily basis may be limited to the first month. Nevertheless, in this particular population these gains may be compromised, because the probability of occurrence of the early selection window appears to be relatively low.

Moreover, the aforementioned conclusions were derived from a liability model, so consideration is required when they are used to make inferences on the observed scale. However, when a REML RR model was applied on the same dataset on the observed scale, the search for periods for which indirect selection appeared to be more effective than direct selection pointed to similar laying stages as the daily liability model. This consistency between models may suggest that results from the liability model could be extrapolated to the observed scale.

In this study the partial production was investigated as an alternative trait for indirect selection, but all periods were examined on the same basis. Alternatively, the use of economic weights in relation to the timing of laying may constitute a more objective-oriented approach to allow focusing on the most interesting periods in terms of genetic information. Breeders may capitalise upon this information in order to maximise the output of selection.

This study has illustrated that the application of a longitudinal threshold model for the study of the daily egg production in turkeys is feasible. The treatment of the daily laying as a binary character made possible the genetic analysis, whereas this was not possible with a more conventional approach on the observed scale. The cost of obtaining a detailed modelling of the genetic covariance structure over time is the increased requirements in computing resources. Nevertheless, breeders can obtain a useful tool to maximise the potential of selection, optimise the use of resources and perform reliable and rapid screening of the output of alternative breeding strategies.

Chapter 7

Bivariate Analysis of Monthly Egg Production and Body Weight in Turkeys Using a Random Regression Model

7.1 Introduction

The current trend in turkey breeding is to select for both growth and reproduction traits. However, it has been shown in Chapter 3 that the correlation between body weight and total egg production was strongly negative in the current population, hindering the simultaneous selection for these characters. Results from a monthly (Chapter 5) and a daily (Chapter 6) model of the egg production have illustrated that heritability changes over the duration of the laying period. This change is partly due to environmental factors and partly due to genetic variance varying over time, implying that gene effects can also have a time-dependent property. This adds a new dimension to genetic correlation between longitudinal traits: Instead of considering it as a fixed measure of the relationship between characters, it may also be treated as time-dependent estimate.

Therefore, a combined analysis of body weight and longitudinal egg production would allow investigation of the structure of the genetic covariance between the characters. The motive for such an analysis is to detect laying stages in which the magnitude of correlation appears to be more favorable than the correlation for total egg production.

As discussed in the previous chapters, random regression (RR) models can provide a robust framework to study laying over time for a univariate analysis. Therefore, the same approach can be used to implement a bivariate analysis of egg production and body weight on a longitudinal basis in order to detect periods for which a breeding opportunity could emerge. An additional benefit of a multi-variate analysis is that it accounts for previous selection and thus, reduces the bias in estimates of variance components (Mrode, 2005). This is of particular relevance to poultry breeding, where birds are preselected using their body weight as a criterion and only qualified birds are selected for egg production. Moreover, this breeding strategy has been performed intensely over numerous generations.

The bivariate analysis employing RR models has been applied to dairy cattle (Gallo *et al.*, 2001; Veerkamp *et al.*, 2001), but no relevant studies exist in poultry breeding. Jaffrezic *et al.* (2004) highlighted the potential of this approach for the combined genetic analysis, but based on simulation data, suggested that the choice of the model relies on the covariance structure of records contained in the examined dataset. In dairy cattle, it has been shown that it is feasible to apply a bivariate RR model to investigate the longitudinal relationship between milking traits (Karacaören *et al.*, 2006).

So, the goal of this study was to investigate the covariance structure between monthly egg production and body weight at the age of fourteen weeks over time. For this purpose, a bivariate model was developed that included fixed and random regression terms for the longitudinal trait (egg production) and a Gaussian trait (body weight). The estimation of genetic parameters aimed to assess the dynamic character of the genetic correlation between the two traits and ultimately to detect periods in which the strength of the genetic association between the traits was weaker compared to that between total egg number and body weight.

7.2 Materials and methods

7.2.1 Traits considered

The traits included in this study were the monthly egg production and the body weight at a specific age. The first trait was identical to that used in the analysis developed in Chapter 5 and consisted of the cumulative egg production of four laying weeks (EGG). Given that records for twenty weeks were available, in total five periods were generated, which corresponded approximately to the monthly egg production. In total, 1920 hens were included in the analysis, originating from 4 generations.

The second trait was the body weight at fourteen weeks (BW14), being the same trait as used in Chapter 3. The primary selection for growth traits was performed at this age, and hence, the number of records was the maximum, compared to weights at later stages for which weight measurements were also available. Moreover, at this time point the selection intensity was also the highest. Finally, it has been shown in Chapter 3 that the genetic correlation between body weights at this age and later stages was very strong structure of records contained in the examined dataset. In dairy cattle, it has been shown that it is feasible to apply a bivariate RR model to investigate the longitudinal relationship between milking traits (Karacaören *et al.*, 2006).

So, the goal of this study was to investigate the covariance structure between monthly egg production and body weight at the age of fourteen weeks over time. For this purpose, a bivariate model was developed that included fixed and random regression terms for the longitudinal trait (egg production) and a Gaussian trait (body weight). The estimation of genetic parameters aimed to assess the dynamic character of the genetic correlation between the two traits and ultimately to detect periods in which the strength of the genetic association between the traits was weaker compared to that between total egg number and body weight.

7.2 Materials and methods

7.2.1 Traits considered

The traits included in this study were the monthly egg production and the body weight at a specific age. The first trait was identical to that used in the analysis developed in Chapter 5 and consisted of the cumulative egg production of four laying weeks (EGG). Given that records for twenty weeks were available, in total five periods were generated, which corresponded approximately to the monthly egg production. In total, 1920 hens were included in the analysis, originating from 4 generations.

The second trait was the body weight at fourteen weeks (BW14), being the same trait as used in Chapter 3. The primary selection for growth traits was performed at this age, and hence, the number of records was the maximum, compared to weights at later stages for which weight measurements were also available. Moreover, at this time point the selection intensity was also the highest. Finally, it has been shown in Chapter 3 that the genetic correlation between body weights at this age and later stages was very strong (close to unity), and so estimates based on BW14 would be also reliable predictors for the other body weight traits. All this suggested that BW14 was the single trait with the greatest potential to reduce selection bias, and therefore, it was selected to be analysed simultaneously with EGG. In total, 20784 birds, both females and males, were included in the analysis from 4 generations.

7.2.2 Genetic model

For the egg production the model included a hyper-factor to account for the fixed effects, including all combinations of hatch, pen and year (see Chapters 5 & 6). The model also included a fixed regression term, derived from the Ali-Schaeffer equation (Ali and Schaeffer, 1987), in order to model the phenotypic trajectory for the mean egg production over the different periods. Two random effects were considered for egg production: one for the additive genetic and one for the permanent environmental component. Legendre polynomials were fitted for the RR. The polynomials were of second order, a choice based on the results presented in Chapter 5.

Hence, the model fitted for the egg production y of hen i at time t was:

$$y_{ijt} = YHP_j + FR_t + \sum_{m=0}^{2} a_{jm}\phi_m(w_t) + \sum_{m=0}^{2} c_{jm}\phi_m(w_t) + e_{ijt} \tag{7.1}$$

where YHP_j was the j^{th} combined fixed effect for the hyper factor and FR_t was the fixed regression terms for time t provided by Ali-Schaeffer equation (Equation 5.1). Terms a_{im} and c_{jm} (with $a_{im} \sim N(0, \sigma_a^2\mathbf{A})$ and $c_{im} \sim N(0, \sigma_c^2\mathbf{I})$; **A** being the relationship matrix between animals and **I** the identity matrix) corresponded to the random effects of the additive genetic and permanent environment effect, respectively, for the animal j for Legendre polynomials $\phi_m(w_t)$ of order m, where w_t was the standardised day t (Kirkpatrick *et al.*, 1990). Finally, e_{ijt} ($e_{ijt} \sim N(0, \sigma_e^2\mathbf{I})$) which was assumed to be homogeneous across all periods.

For the body weight the fixed effects consisted of the following: sex, hatch, cohort and an interaction term between sex and cohort, in accordance to the model used earlier (see

Chapter 3). The model for body weight included also a dam random effect, which was treated as a environmental one (i.e. cov(a,d) = 0).

Hence the model fitted for the body weight y for individual i of dam m was

$$y_{ijklm} = S_j + H_k + C_l + S.C_{jl} + a_i + d_m + e_{ijklm} \tag{7.2}$$

where S_j was the j^{th} sex, H_k the k^{th} hatch, C_l the l^{th} cohort and $S.C_{jl}$ the interaction between the j^{th} sex and l^{th} cohort. Terms a_i and d_m (with $a_i \sim N(0, \sigma_a^2 \mathbf{A})$ and $d_m \sim N(0, \sigma_d^2 \mathbf{I})$) respectively; $\mathbf{A}$ being the relationship matrix between animals and $\mathbf{I}$ the identity matrix) corresponded to the random effects of additive genetic effect of individual i and the permanent environmental effect of its dam m. Finally, e_{ijklm} was the residual term ($e_{ijklm} \sim N(0, \sigma_e^2 \mathbf{I})$).

As derived from 7.1 and 7.2, the model of the monthly egg production included a second random effect, corresponding to the permanent environment of the hen, whilst this effect cannot be fitted to body weight model, since only one record was available per bird. This inconsistency would result in covariance matrices for the two random effects having unequal dimensions, and so it would not be possible to obtain a total covariance matrix, corresponding to phenotypic variance. To overcome this constraint, body weight data points were doubled by adding and subtracting from the original records a small quantity (e.g. 0.02). This was equivalent as having two measurements for body weight with a small error. This treatment allowed to obtain symmetric covariance matrices having the same dimensions for both random effects and estimate genetic correlations between the two traits.

In conclusion, six traits were included in the analysis, 5 corresponding to the time points available for egg production and one for BW14. Therefore, all covariance matrices had the same dimensions (6×6). However, due to the use of RR, the number of variance components estimates reduced to 10 unique items that formed a 4×4 coefficient matrix (**K**) for both the genetic and permanent environmental random effect. The structure of matrix **K** is given in 7.3. There it is shown that the diagonal elements accommodate the variance of body weight (BW) and the three elements of the RR, i.e the intercept (int), the linear (lin) and square (sq) regression coefficients. The off-diagonal elements correspond to the covariances between the diagonal elements.

$$K = \begin{bmatrix} \sigma^2_{BW} & \sigma_{BW,int} & \sigma_{BW,lin} & \sigma_{BW,sq} \\ \sigma_{BW,int} & \sigma^2_{int} & \sigma_{int,lin} & \sigma_{int,sq} \\ \sigma_{BW,lin} & \sigma_{int,lin} & \sigma^2_{lin} & \sigma_{lin,sq} \\ \sigma_{BW,sq} & \sigma_{int,sq} & \sigma_{lin,sq} & \sigma^2_{sq} \end{bmatrix} \tag{7.3}$$

7.2.3 Estimation of genetic parameters

The covariance matrix for the genetic effect **G** was given as the product of

$$\mathbf{G} = \Phi^T \mathbf{K_g} \Phi \tag{7.4}$$

where Φ was the 4×6 matrix of the time covariates for the Legendre polynomials and $\mathbf{K_g}$ the 4×4 matrix containing the estimate of variance for BW14 and the intercept and the two random regression coefficients for EGG. Likewise, the covariance matrix $\mathbf{K_{pe}}$ for the permanent environmental effect **PE** was computed. Finally, the residual covariance matrix **R** was constructed from a 6×6 identity matrix, having the first diagonal element equal to the residual variance for BW14 and the rest five diagonal elements equal to the homogeneous residual variance for EGG. Finally, the phenotypic covariance matrix **P** was computed as the sum of **G**, **PE** and **R**.

The heritability (h_i^2) for period i and the genetic correlation (r_{ij}) between periods i and j were estimated as follows:

$$h_i^2 = \frac{G_{i,i}}{Pi,i} \tag{7.5}$$

$$r_{ij} = \frac{G_{i,j}}{\sqrt{G_{i,i} \times G_{j,j}}} \tag{7.6}$$

by substituting the appropriate elements of **G** and **P** matrices.

The approximate standard errors (ASEs) of genetic parameters were estimated using the inverse of the average information matrix. The dimensions of this matrix were 22×22, since every line corresponded to the unique covariance between each variance component estimate (10 for additive genetic effects, 10 for permanent environment and 2 for residual) and all the other ones. The method proposed by Fischer *et al.* (2004) was extended in order to estimate not only ASEs for the heritability, but also for the correlation estimates between all traits.

According to Fischer *et al.* (2004), equation 7.4 can be also written using vector forms of the matrices (i.e. by stacking the columns of matrices to form a vector). Because all the matrices involved are symmetric, it is obvious that the off-diagonal elements are contained twice in the vector form. In order to compute the covariance of the covariance matrix, the output of ASReml (file *.vvp) can be used. In this file, only the covariance of the unique variance components is contained. In this particular example, as explained above, 10 unique elements are contained in $\mathbf{K_g}$. Therefore, the challenge was to devise a formula to convert the 10×10 matrix from the ASReml output to the vector form of the covariance of covariance matrix of the regression coefficients for the additive genetic effect Vec($\mathbf{K_g}$). The dimensions of this new matrix were 16×16 (i.e. $(order)^2 \times (order)^2$). The new matrix then can be used to estimate the vector form of the matrix containing the covariance of covariance matrix Vec(**VG**), as follows:

$$Vec(\mathbf{VG}) = (\Phi \otimes \Phi)' Vec(\mathbf{K_g})(\Phi \otimes \Phi) \tag{7.7}$$

The dimensions of **VG** were equal to 36×36 (i.e. $(traits)^2 \times (traits)^2$). The same procedure was followed to estimate respective matrices for $\mathbf{K_{pe}}$. This matrix contains redundant elements, therefore it can be summarised to a 6×6 matrix, which contains the variance of the corresponding element of $\mathbf{K_g}$. The mapping function relied on the number of traits and it was arbitrarily expressed for the specific problem. As a guideline, the following rules were used to define the diagonal and non-diagonal elements:

$$\begin{gathered}
\mathrm{VG}_{i,i} = \left(\frac{G_{i,i}}{P_{i,i}}\right)^2 \left(\frac{VG_{6i,6i}}{G_{i,i}^2} + \frac{VP_{6i,6i}}{P_{i,i}^2} + \frac{2covGP_{6i,6i}}{G_{i,i}^2 P_{i,i}^2}\right) \\
\mathrm{VG}_{i,j} = r_{ij}^2 \left(\frac{VG_{6i,6i}}{4G_{i,i}^2} + \frac{VG_{6j,6j}}{4G_{j,j}^2} + \frac{VG_{6i,6j}}{G_{i,j}^2} + \frac{VG_{6i,6i}VG_{6j,6j}}{2G_{i,i}G_{j,j}} + \frac{VG_{6j,6j}VG_{6j,6j}}{G_{i,i}G_{i,j}} + \frac{VG_{6i,6i}}{G_{i,j}G_{j,j}}\right) \quad (7.8)
\end{gathered}$$

where r_{ij}^2 is the square of the correlation coefficient between periods i and j, as estimated in 7.6. The functions in 7.8 were formulated according to the standard formulas for obtaining the variance of variance components (Gilmour *et al.* (2001); see also Equation 4.11). The diagonal elements of **VG** corresponded to the variance of the heritability and the off-diagonal to the variance of the correlation between the i^{th} and j^{th} trait. ASEs were obtained as the square root of the appropriate elements of **VG**.

The latest version of ASReml (v.2) was used to estimate the variance components. The translation of the estimates of regression coefficients into covariance matrices over time was performed using the custom-developed program. Routines for the estimation of ASEs for RR were included in this program, since there was no available software to compute them. The validity of the ASEs estimates was tested by repeating the estimation of SEs using the data for RR model as in Chapter 5.

7.3 Results

The results for genetic, environmental and total phenotypic covariance over time for the monthly egg production are summarised in Tables 7.1, 7.2 and 7.3 respectively. There it is shown that genetic variance was increased at the onset of laying, significantly reduced at the peak of production, in second month, and towards to the end of the production period was increased again. In contrast, the permanent environmental component was monotonically increased over time. The same trend applied also for the total phenotypic variance.

Heritability was estimated for egg production that corresponded to each individual period of laying and also for the cumulative yield from onset and up to the fifth period. Since the cumulative yield of all periods is equivalent to the total egg production, the heritability of the total egg number was found to be 0.14. These estimates are presented in Table 7.4. The change in heritability estimates of egg yield for each individual period is also visualised in Figure 7.1.

Table 7.1: Genetic (co)variance over time for monthly egg production.

Period	1	2	3	4	5
1	**3.16**	1.26	0.42	0.63	1.88
2		**1.11**	1.11	1.26	1.59
3			**1.58**	1.81	1.83
4				**2.27**	2.64
5					**4.00**

Table 7.2: Permanent environmental (co)variance over time for monthly egg production.

Period	1	2	3	4	5
1	**7.98**	8.05	7.58	6.59	4.98
2		**10.81**	12.02	11.69	9.83
3			**14.51**	15.06	13.69
4				**16.68**	16.47
5					**18.26**

Table 7.3: Total phenotypic (co)variance over time for monthly egg production

Period	1	2	3	4	5
1	**19.37**	9.31	8.01	7.18	6.86
2		**20.52**	13.13	12.96	11.41
3			**24.69**	16.88	15.51
4				**27.73**	19.11
5					**30.87**

The heritability for body weight at fourteen weeks was estimated to be equal to 0.31±0.03. For the same trait, the environmental maternal effect accounted for the 6% of the total phenotypic variance, suggesting a considerable contribution of dam component, being

Table 7.4: Heritability estimates for the egg production of individual periods (left column) and cumulative production up to period *x* (right column).

Period	h^2 period	h^2 cumulative
1	0.20 ± 0.04	0.20 ± 0.04
2	0.07 ± 0.02	0.15 ± 0.03
3	0.07 ± 0.02	0.12 ± 0.02
4	0.09 ± 0.03	0.12 ± 0.03
5	0.15 ± 0.03	0.14 ± 0.03

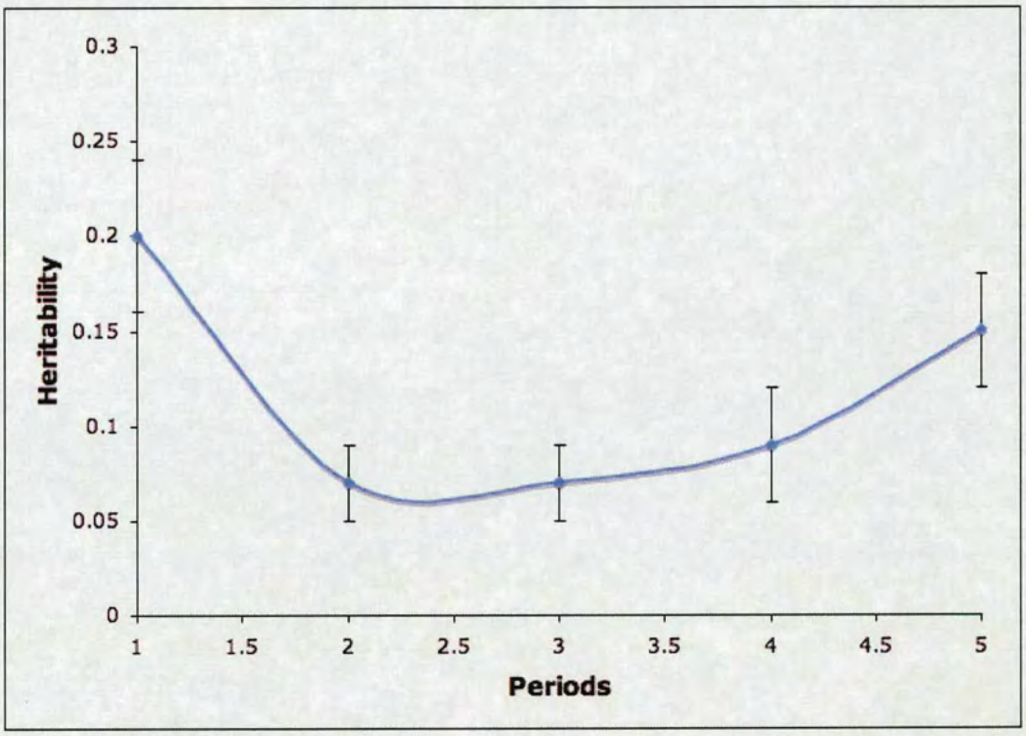


Figure 7.1: Heritability profile for all periods.

consistent with the conclusions in Chapter 3.

The estimates of phenotypic and genetic correlation are summarised in Table 7.5. Overall, the general pattern of the genetic correlation between laying stages was consistent with the findings from the previous chapters. In particular, the genetic correlation between consecutive periods was found to be strong, while a weak association appeared to link the onset and the middle of lay.

The phenotypic correlation between all periods of laying and body weight was weak, yet statistically different from zero. In contrast, the genetic correlation between laying stages and body weight was strongly negative. Periods three and four, which corresponded to the middle part of the whole production period, had the lowest estimates with body weight among all the laying subperiods. Despite the fact that these estimates appeared to be considerably lower than the respective estimates of body weight with the rest of the periods, due to their large standard error, it was not not possible to reject the null hypothesis that the genetic covariance changed over time. Moreover, these estimates could not support the hypothesis that a specific laying stage had a weaker genetic correlation with body weight than its correlation with total egg production.

Table 7.5: Phenotypic (above the diagonal - in italics) and genetic correlation (below the diagonal - regular fonts) betweem body weight at 14 weeks (BW14) and egg production of individual periods (P1 to P5).

	BW	**P1**	**P2**	**P3**	**P4**	**P5**
BW	1	*-0.20±0.03*	*-0.12±0.03*	*-0.08±0.03*	*-0.10±0.03*	*-0.16±0.03*
P1	-0.57±0.11	1	*0.46±0.02*	*0.36±0.02*	*0.31±0.02*	*0.27±0.02*
P2	-0.47±0.17	0.67±0.12	1	*0.58±0.02*	*0.54±0.02*	*0.45±0.02*
P3	-0.29±0.18	0.19±0.21	0.84±0.07	1	*0.64±0.02*	*0.56±0.02*
P4	-0.41±0.14	0.23±0.19	0.80±0.10	0.95±0.03	1	*0.66±0.02*
P5	-0.63±0.11	0.59±0.14	0.75±0.14	0.73±0.12	0.88±0.06	1

7.4 Discussion

The application of the bivariate model for body weight and monthly egg production, modelled by random regression terms, offered the benefit of accounting for on-going selection on the population for growth and reproduction traits. The heritability estimates for egg production, whilst being consistent with previous results in pattern, they tended to be greater in magnitude. Although there was some indication that periods in the middle of the laying period had a weaker association with body weight, there was not enough statistical evidence to establish such an advantage. Nevertheless, a genetic evaluation employing bivariate RR models may still offer a benefit, in terms of the increased heritability estimates for the egg production.

The population studied in this analysis was selected for both growth and reproduction traits. This was not accounted for in the previous chapters, since a univariate longitudinal analysis of egg production had been performed. This may have resulted in underestimated heritabilities. The application of a bivariate model can reduce the bias on the estimation of variance components because it includes information on the on-going selection. Moreover, a higher precision of the estimates is achieved (Mrode, 2005).

There are benefits from the BLUP analysis of two or more strongly correlated traits, such as that an increase in the accuracy of evaluations is expected (Thompson and Meyer, 1986). Moreover, a notable benefit for the low heritable trait is that the heritability estimate is increased, when analysed simultaneously with a strongly correlated, but more heritable trait (Mrode, 2005). In situations where a culling decision on animals has been made on early stages based on the performance on one trait and only selected individuals have records for the other trait, then a multivariate analysis removes the bias that a univariate analysis of the second trait would impose. These advantages of the multi-trait genetic evaluation are very relevant to the current population, because body weight and egg production were highly correlated and selection was performed on the former previous to selection for the latter.

The heritability estimate for the body weight was similar to the estimate from the multivariate analysis presented in Chapter 3 (0.37±0.03) and in good agreement with literature (Buss, 1990). The small difference between current and previous estimate can be explained by the fact that the latter analysis included also weight at 19 and 24 weeks. Although primary selection was performed at 14 weeks, a second selection stage was taking place at 24 weeks. Therefore, the increased heritabilty estimate from the previous chapter may be partly explained by the fact that this model contained more information on the selection history of the population. Maternal effects were found similar in magnitude to the previous estimates, providing additional evidence on the importance of including them in the genetic analysis of growth traits, as this was discussed in Chapter 3.

The pattern of estimates of additive and permanent environmental variance for monthly egg production was comparable to that from the univariate RR model, as this was reflected in the similar shape of heritability profiles from the single- and two-trait analysis. Nevertheless, the estimates from the bivariate model were systematically increased compared to the univariate ones, due to the reasons explained above. Nevertheless, this difference was not statistically significant, due to the relatively large standard error of the estimates (Table 7.6).

Table 7.6: Comparison of heritability estimates with their standard errors for a bivariate (2^{nd} order) and two univariate RR models (2^{nd} and 3^{rd} order respectively)

Period	Bivariate	RR2	RR3
1	0.20 (0.04)	0.12 (0.03)	0.13 (0.04)
2	0.07 (0.02)	0.05 (0.02)	0.11 (0.03)
3	0.07 (0.02)	0.07 (0.03)	0.08 (0.03)
4	0.09 (0.03)	0.07 (0.03)	0.05 (0.03)
5	0.15 (0.03)	0.07 (0.03)	0.08 (0.03)

The benefit from the multivariate analysis, in terms of increased heritability for egg production, is of particular interest in case of the longitudinal analysis of laying. It has been previously shown that these models can provide the genetic covariance structure over time, but their benefit was hampered by the fact that it was not possible to account for the on-going selection for other traits. However, this study has shown that it is possible to analyse simultaneously a dynamic and a static trait and capitalise upon the increased accuracy of estimates. This improvement can be considerable, as our results have suggested. When a univariate RR model was applied, the cumulative heritability over all periods, which corresponded to total production, it was found to be 0.07. In contrast, in this analysis, the overall cumulative heritability was estimated to be 0.14.

An alternative approach for accounting for both the longitudinal aspect of laying and selection on body weight would be to employ a multi-trait model. However, in this case the high dimensionality of the problem space would impose considerable constrains on the solution convergence. So, were such a model to be used with the present dataset, the amount of covariances estimated would be equal to 43 (21 for additive genetic, 1 for maternal effects and 21 for residuals) instead of 23 in a RR model(10 for additive genetic, 10 for permanent environmental, 1 for maternal effects and 2 for residuals). Thus, the advantage of RR models, as already been shown in Chapter 5, becomes more obvious in the context of the current study.

There was not sufficient evidence to confirm the initial hypothesis for a varying over time genetic covariance between laying stages and body weight, due to the large ASEs of correlation estimates. The precision of the estimates depends on the number of records available and how individuals are connected through the relationship matrix. Since no additional observations were available for analysis, a more detailed insight of the longitudinal changes of genetic covariance was obtained by monitoring the proportion of genetic variance of egg production that is not associated with body weight. The proportion of genetic variance of egg production σ^2_{unex} that is not explained by body weight is given from the following formula:

$$\sigma^2_{unex} = \sigma^2_a(1 - r^2) \tag{7.9}$$

where, r^2 is the square of the correlation coefficient. From 7.9, the proportion of genetic variance that is not explained by the regression on body weight was very high for both the middle periods (91% and 83% repsepctively for periods 3 and 4). Nevertheless, these periods had also the lowest heritability estimates. Hence, even if it could be suggested that the genetic correlation relaxed for these periods, still they would not represent a breeding opportunity, because of their low heritability. The use of either partial production from period 3 or 4 is less efficient compared to the direct selection for total egg production. In conclusion, even if genetic covariance was shown to be variable over time from another study using more data that would achieve higher estimate precision, it is expected that the favourable periods could not be exploited for making simultaneous selection of the antagonistic traits more efficient than the use of total egg production and body weight.

This study showed that the application of bivariate analysis offered increased heritability estimates of the monthly egg production over previous estimates obtained from a univariate longitudinal analysis. However, it was not possible to pinpoint a laying stage for which the selection for partial records would be at least as efficient as the selection for total egg number, whilst the correlation with body weight being lower. This suggests that the number of options for breeders to manage the unfavourable strong association between those two traits is reduced.

Chapter 8

General Discussion

8.1 Objectives revisited

The first indication of investigating the reproduction physiology of avian species dates back to Aristotle[1]. Since then, a substantial number of key mechanisms involved in egg laying have been unravelled. These suggest that a complex regulatory network, heavily relying on the endocrine system, governs the phenotypic expression of the trait perceived as egg production. Nevertheless, little effort has been put into translating these findings in a quantitative genetics context. From a breeder's point of view, the objective is to maximise the total number of eggs laid within a specified period. Although the rationale for using this trait derives from its economic importance, it implicitly ignores the longitudinal aspect of egg laying.

Therefore, the main objective of this thesis was to contribute to the investigation of egg laying genetics over time. The motive was to explore whether considering the egg production as a dynamic trait would offer additional benefits over treating it as a static character. These advantages may refer to both a theoretical and applied basis. In the former context, the development of models could enhance our understanding of egg laying and indicate new research paths for more detailed studies. In the latter context, these models may provide more accurate description of egg laying, identify new traits as an alternative to the total egg number and thus, improve the output of the breeding programs performed on commercial populations.

Different aspects of egg production over time have been studied in the previous chapters. Although the results have been discussed in the respective sections, here they are linked together, forming a wider picture. Key findings from each chapter are summarised first. This review then considers their implications and thus combines them in an holistic overview of the longitudinal character of egg laying under a quantitative genetics perspective.

---

[1] *Aristotle, Historia Animalium VI. III, 561a4-21 - First description of chicken embryo development.*

8.2 Key findings

The main findings of the five research chapters are summarised below.

Chapter 3: The objective of this chapter was to investigate the strength of the genetic association between growth and reproduction traits in turkeys selected for body weight, conformation and egg production. Two distinct populations but derived from the same heavy turkey female line and situated in different locations (UK and USA), were used to estimate genetic parameters using multivariate REML for the following traits: body weight at 14 (BW14), 19 (BW19) and 24 (BW24) weeks of age and total egg number (EGG). A Box-Cox transformation was applied to egg production data to reduce the impact of non-normality. The heritability estimates for each trait for the UK and USA populations respectively were: BW14 0.37 and 0.48; BW19 0.34 and 0.43; BW24 0.28 and 0.43; EGG 0.22 and 0.34. The genetic correlation between body weight at all ages and total egg production was strongly negative, reaching a value of -0.75 for the UK and -0.55 for the USA population. The comparison of these results with published estimates in turkeys led to the hypothesis that the genetic correlation may get stronger in magnitude following selection for increased body weight. This could arise from fixation during selection of genes favouring larger weights but with minimal effect on egg production, leaving the segregating genetic variation dominated by pleiotropic loci with antagonistic effects on the traits studied. Thus, in order to avoid continued selection for body weight reducing egg production, alternative selection strategies should be considered.

Chapter 4: An alternative approach was applied for the longitudinal study of egg production. A time-to-event trait was formulated that consisted of the time required for a hen to lay 82 eggs, which was the grand average of the egg production over the five generations available. This definition of the terminating event allowed the fitting of a Weibull survival model that included a random effect, corresponding to the sire component. This new trait accounted for features of laying patterns, such as the rate and persistency of lay, yet it was strongly correlated with the total egg number. This suggested that the time-to-event trait can be considered as a selection criterion for the improvement of the egg production.

Supportive arguments are first, that the normality assumption that hampers the estimation of genetic parameters is no longer an issue under a survival analysis context. Second, an explicit censoring rule can be suggested for excluding outliers. Third, the genetic variance available appears to be increased, as implied by the relatively high heritability estimate, derived from a sire model. Nevertheless, for the purposes of genetic evaluation, survival models are still under development with the most notable constraints being the inability to obtain BLUP estimates of breeding values and to perform multivariate analysis. In conclusion, this study showed that the application of survival analysis is feasible and has the potential to allow breeders to summarise effectively data for improving egg production, as research is performed to resolve issues that limit the scope of its application.

Chapter 5: Random regression (RR) models have become a popular methodology for the genetic study of longitudinal data in dairy cattle, however their implementation in poultry breeding is limited. Therefore, the first objective of this chapter was to investigate the application of RR models for the genetic analysis of egg production in turkeys. Data collected from a heavy female line were used to estimate genetic parameters with two RR models, one having second order Legendre polynomials as regression over time (RR2) and another with third order polynomials (RR3). The second objective was to benchmark the performance of RR models with more conventional methods; so, genetic parameters were re-estimated using a multi-trait (MT) and a repeatability (REP) model. In order to assess the model efficiency at predicting missing values two reduced datasets were used. The first one included a number of deleted records evenly spread over all periods. For each model the predicted values of those deleted records were compared with the true values. The second dataset was generated by eliminating the last period. Thus, RR models were further compared against each other by estimating the mean square error of the predictions for the missing period. In the first test, the REP model had the poorest performance in predicting missing values. Heritability estimates from the RR2 and MT models were close to each other, while the RR3 estimates were different. Both RR models demonstrated better predictive ability than the MT model. However, when only RR models were compared in the second test, the RR2 resulted in the smallest mean square error. This provided evidence that the RR3 model overfitted the data, suggesting that the choice of the appropriate polynomial order requires careful consideration. This study has illustrated that the application of RR models for the genetic analysis of the egg production in turkeys is not only feasible, but also offers higher accuracy of prediction than alternative longitudinal models.

Chapter 6: The previous chapter indicated that the longitudinal analysis of egg production can benefit from the application of RR models. In order to study in detail the laying patterns, the length of individual periods included in RR model had to be reduced. However, the model fitted failed to converge, when periods shorter than four weeks were fitted. Paradoxically, the extremely short interval of a single day was more tractable by treating the daily egg production as a binary character. Therefore, a threshold liability model was used, which included RR terms for both a residual hen and a sire genetic effect,implemented under a Bayesian framework. The objectives of this chapter were, first, to estimate genetic parameters for daily egg production over the whole laying period; and second, to explore the possibility of using the information obtained to investigate alternative selection strategies for the genetic improvement of laying turkeys. The shape of the heritability profile and the correlation patterns were in accordance with the estimates from the monthly RR model of the previous chapter, albeit they provided a more detailed insight into laying dynamics. The efficiency of selection using the cumulative production of subperiods was compared to selection for total egg number. The periods for which the indirect selection was more efficient constituted favourable breeding opportunities. Such efficient selection windows were observed for both the first and the last month of production, with the latter appearing to be more favourable due to the higher genetic gains associated. Although the potential advantage of using partial records needs to be put into context along with other factors that govern the design of efficient breeding programs, it might offer a tool for improving egg production in a multi-trait selection scheme. Since results referred to the liability scale, an approximation was used in order to translate them to the observed scale.

Chapter 7: In previous chapters it was demonstrated that the genetic variance changes over time, implying that the genetic effects on egg production vary during the laying period. Hence, it can be postulated that the genetic association between loci controlling body weight at a certain age and loci affecting the egg production over specific laying stages may also be varying in magnitude. Therefore, the objective of this chapter was to investigate the genetic covariance structure on a longitudinal basis in order to indicate periods in which the genetic correlation appears to be more favourable under the currently applied breeding program. For this purpose a bivariate RR model was formulated that included body weight at 14 weeks (BW14) and monthly egg production, as specified in Chapter 5. Although the pattern of the heritability estimates was consistent with the results from the previous chapters, the magnitude of the heritability for egg production was increased. This happened because the bivariate model accounted for the selection for body weight over many generations and removed the bias in the genetic parameter estimates of the egg production from the univariate RR model. The genetic correlation patterns provided evidence that the genetic association between the two traits tended to vary over time, but the differences detected were not significant due to the large standard errors of the estimates. Nevertheless, even if the benefit of these laying stages in terms of reducing the strength of the association was significant, this would not constitute a breeding opportunity. The reason is that the periods in which genetic correlations tended to be weaker were also the periods with the lowest genetic variance. So the potential benefit from the reduced correlation appeared to be counterbalanced by the loss of genetic gains, due to the reduced heritability. This implies that meticulous evaluation of the candidate population is warranted before the bivariate RR model is deployed in large-scale breeding programs.

8.3 Implications

The longitudinal genetic analysis of egg production revealed time-depended characteristics of genetic variance that potentially can assist in unraveling the genetics of laying over time. Ultimately, this may be exploited by the design of more efficient breeding programs. In the next sections, these implications are discussed in detail.

8.3.1 Laying dynamics

Among the key findings of this thesis was that the genetic variance of egg production has a dynamic character. In order to illustrate this point further, the patterns of inheritance can be analysed by estimating the eigenvalues of the regression coefficients of RR models and plotting the corresponding eigenfunctions over time (Kirkpatrick *et al.*, 1990). In this context, the components of genetic variance can be clustered and thus, each eigenvalue corresponds to a linear combination of a set of genetic factors. The decomposition of the genetic variance therefore enables the assessment of how the contribution of each factor set in the total genetic variance changes over time[2]. The eigenvectors were estimated from the genetic covariance matrix obtained from the daily model (matrix V in Equation 6.3). The rationale for choosing the daily model was that it not only provided more time points than the monthly one, but also referred to the underlying scale, which can be considered more suitable for the investigation of heritability of time trends. Since the remaining genetic effects were included in the residual hen effect, the eigenfunctions of the environmental covariance matrix were also estimated.

The first two genetic eigenvalues accounted for approximately 98% of the sum of all genetic eigenvalues (74.5% and 23.3% respectively). The three genetic eigenfunctions were plotted against time (Figure 8.1), where it is seen that all of them change considerably over time. Furthermore, the first and the third eigenfunctions change algebraic sign during the laying period, which implies that the effect of the corresponding genetic components is different between the periods. A possible interpretation is that different sets of genetic factors are activated during the laying period. Moreover, the selection on eigenfunctions associated with contrasting gene effects is expected to vary according to the sign of the variable. For example, selecting for the leading eigenfunction would increase egg production in the beginning and finishing laying stages, whereas it would decrease it in the

---

[2] Additionally, the eigenvalues have been used to test the significance whether the inclusion of higher order polynomials as additional regression terms improves the fit of the RR model or not (Meyer and Hill, 1997). This has been discussed in Chapter 5, where it was stressed that the improvements on the model fit as a consequence of increasing the number of regression coefficients should be counterbalanced by the problem of overfitting the data.

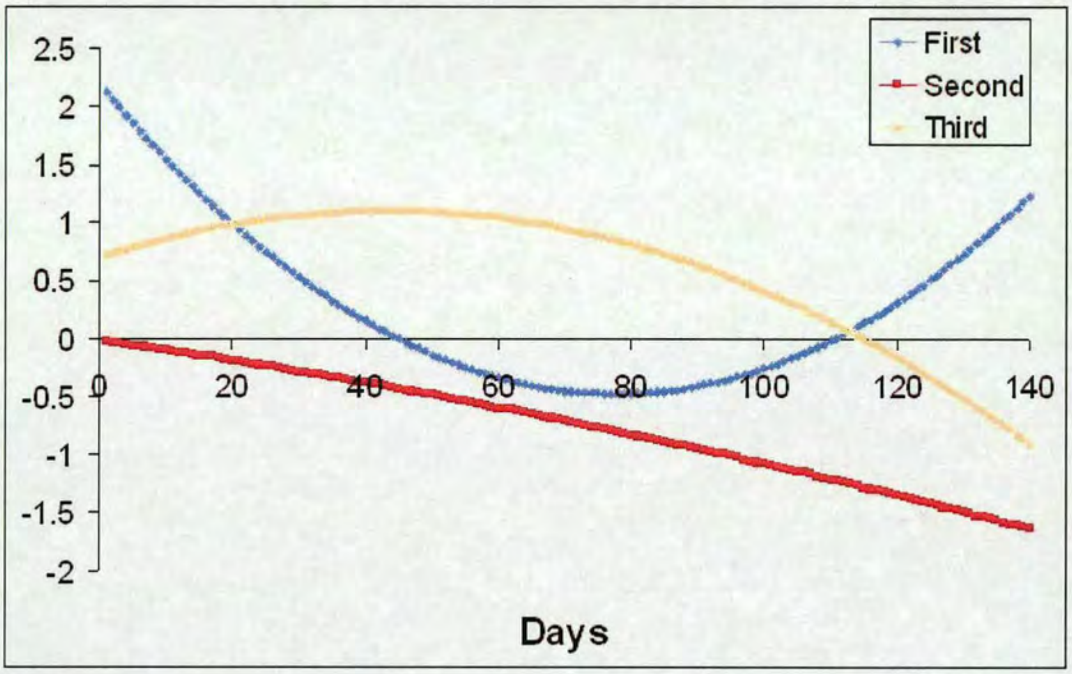


Figure 8.1: Eigenfunctions of the genetic covariance function coefficient matrix.

middle phase.

The contrasting genetic effects over time have been already implied by the genetic correlation surface presented in Figure 6.5. There, it was demonstrated that the beginning and finishing stages of laying were negatively correlated with the middle period, covering the interval between 45 and 110 days. This coincided approximately with the period for which the leading eigenfunction changes sign. Also, day 45 was the day where the heritability started to rise again after reaching its minimum value. At the phenotypic level, the change of sign of the leading eigenfunction coincides with the beginning of the decline in the hen-day egg production after the peak at 39 days, which signifies the upturn of genetic variance (Figure 6.1). The phenotypic correlation was also weak for the periods for which the first eigenfunction was negative. In conclusion, the plots of genetic eigenfunctions and genetic covariance provided complementary evidence to postulate that different sets of genetic factor are activated during the laying period.

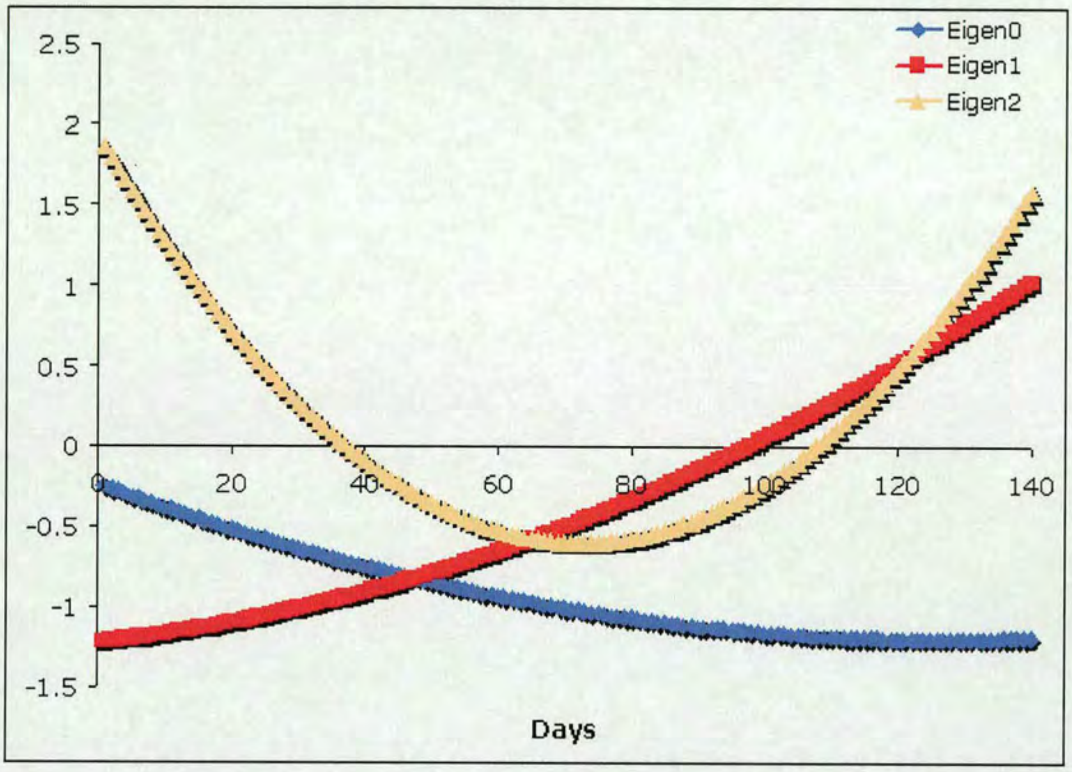


Figure 8.2: Eigenfunctions of the residual hen covariance function coefficient matrix.

The first two eigenvalues of the residual hen effect accounted for approximately 98% of the sum of all eigenvalues (80.7% and 17.4% respectively). In this case, the eigenfunction plot displays a different picture from the genetic plot (Figure 8.2). The first eigenfunction appears to be less variable over time and the second eigenfunction increases constantly over time. Despite its small contribution to the sum of eigenvalues, the pattern of the third eigenfunction may be interesting. It appears to have a very similar shape to the leading genetic eigenfunction. Since all remaining genetic components are included in the residual hen effect, this high similarity provides a tantalising suggestion that the third residual hen eigenfunction is associated with the same genetic factors as in the first genetic eigenfunction.

The eigenfunction profiles presented above appeared substantially different from those for milk production in dairy cattle. There, it was demonstrated that the leading genetic eigenfunction appeared to be approximately constant during the lactation and only the second eigenfunction, which accounted for a smaller portion of the total genetic variance, indicated contrasting gene effects over time (Kirkpatrick *et al.*, 1990; Olori *et al.*, 1999; Togashi and Lin, 2006). A suggestive explanation of this notable difference could be the significance of age-related changes on egg production physiology, because avian species are short-lived and a decline in their reproduction performance is observed toward the end of their production life, including decreased egg production and clutch size [3] (Ottinger, 1992; Wu *et al.*, 2005).

The decomposition of genetic variance, presented in Figures 8.1 and 8.2, was the first performed not only in turkeys, but also for any poultry. This analysis may have implications for unraveling the genetics of egg laying. The variability over time of the genetic eigenfunctions implies that the associated genetic factors affect egg production accordingly to the laying stage. Plausible explanations could be that different genes affect specific laying stages. The results of Minvielle *et al.* (2006) performed in Japanese quail provided support for the existence of QTLs influencing a specific laying phase. There, several QTLs have been detected that had an effect only at a particular laying stage. This also led the authors to postulate the existence of QTLs affecting the shape of laying curve.

8.3.2 Capitalise upon laying dynamics

As discussed above, the dynamic character of egg laying raises many questions for basic studies. Nevertheless, the challenge is also to translate the findings from this thesis to benefits in a quantitative genetics context to enhance breeding strategies.

The simple approach to select for egg production is by counting the number of eggs over a fixed period of time. Although the total egg number implicitly accounts also for some aspects of the dynamics of lay, the analytical focus is placed on the total volume of production. However, in Chapter 4 it was shown that the opposite approach (i.e. counting

---

[3] The decreased clutch size suggests a more erratic laying patterns towards the end of the production period, which was also evident in the data used in this thesis.

time to lay a fixed number of eggs) appeared to be at least as efficient as the total egg-number strategy in terms of the expected genetic progress. Not only was the time-to-event trait strongly correlated to the total egg production, but also accounted for traits related to laying patterns, such as the rate and persistency of lay. These time-related laying traits have not been intensively selected by turkey breeders, implying that there is a considerable genetic variance to be exploited for accelerating genetic progress. The results in Chapter 4 support this hypothesis. Nevertheless, in this thesis apart from investigating the potential of survival analysis, the need has also been stressed for advances in methodological issues that would allow the wide-scale application of this framework in poultry breeding.

The aforementioned analysis indicated that, first, time trends are significant for genetic evaluations, and second, that it is possible to use partial records for genetic evaluations without information loss. A direct benefit from reducing the recording time is the better allocation of resources. However, in order to identify these informative periods it is important to have an insight into the genetic relationship between production at different points of the laying period. In Chapter 5 it has been shown that RR models can provide the machinery to estimate effectively the genetic covariance structure between all laying days.

The RR models provided a detailed model of the genetic covariance structure between all time points of the laying period. This can provide a tool to project partial production over any required number of days and assess the genetic correlation with total production. So, it is possible to target selection only to laying stages where indirect selection for partial records appears to be at least as efficient as direct selection for total egg production.

In Chapter 6 I discussed how the covariance structure over time, derived from a RR model, can be used to assess the efficiency of selection based on partial records. There, it was suggested that the cumulative production of the first month can be used without compromising the expected genetic progress when using full production records. The benefit of an early selection is that it permits breeders to use a hen's own performance for genetic evaluation purposes before the completion of the production cycle. In a routine breeding program, where there is a continuous flow of eggs that are hatched and poults are sent time to lay a fixed number of eggs) appeared to be at least as efficient as the total egg-number strategy in terms of the expected genetic progress. Not only was the time-to-event trait strongly correlated to the total egg production, but also accounted for traits related to laying patterns, such as the rate and persistency of lay. These time-related laying traits have not been intensively selected by turkey breeders, implying that there is a considerable genetic variance to be exploited for accelerating genetic progress. The results in Chapter 4 support this hypothesis. Nevertheless, in this thesis apart from investigating the potential of survival analysis, the need has also been stressed for advances in methodological issues that would allow the wide-scale application of this framework in poultry breeding.

The aforementioned analysis indicated that, first, time trends are significant for genetic evaluations, and second, that it is possible to use partial records for genetic evaluations without information loss. A direct benefit from reducing the recording time is the better allocation of resources. However, in order to identify these informative periods it is important to have an insight into the genetic relationship between production at different points of the laying period. In Chapter 5 it has been shown that RR models can provide the machinery to estimate effectively the genetic covariance structure between all laying days.

The RR models provided a detailed model of the genetic covariance structure between all time points of the laying period. This can provide a tool to project partial production over any required number of days and assess the genetic correlation with total production. So, it is possible to target selection only to laying stages where indirect selection for partial records appears to be at least as efficient as direct selection for total egg production.

In Chapter 6 I discussed how the covariance structure over time, derived from a RR model, can be used to assess the efficiency of selection based on partial records. There, it was suggested that the cumulative production of the first month can be used without compromising the expected genetic progress when using full production records. The benefit of an early selection is that it permits breeders to use a hen's own performance for genetic evaluation purposes before the completion of the production cycle. In a routine breeding program, where there is a continuous flow of eggs that are hatched and poults are sent

Table 8.1: Heritability estimates (h^2) for the egg production in the first month of the laying period derived from a univariate animal model (h^2 with standard error (SE) for untransformed and Box-Cox transformed values) and threshold RR model (h^2 with standard deviation (SD) of the posterior distribution are given both in liability and observed scale).

Model	**Comments**	**Heritability**	**SE/SD**
Animal model	Untransformed records	0.09	0.03
Animal model	Transformed records	0.20	0.04
Threshold RR model	Observed scale	0.30	0.08
Threshold RR model	Liability scale	0.42	0.09

to rearing farms while their dams continue laying eggs, breeders will be able to combine individual partial performance with parent information, increasing the accuracy of prediction of breeding values, which in turn would accelerate the genetic progress achieved per generation.

Nevertheless, in a scenario where the partial production of the first month was to be used, the magnitude of the heritability would also be significant for determining the genetic gains. In order to contrast the estimates derived from the daily RR model, heritability was also estimated from a univariate model, similar to this used in Chapter 3, for untransformed and Box-Cox transformed records of the first month of production. The results are summarised in Table 8.1. There, it appears that higher heritability estimates were obtained when a RR model was employed.

Nevertheless, caution in warranted when selection pressure is put on the production of the first month (Flock, 1977). The increased genetic variability that is detected in this period is probably due to the differences among birds with regards to their sexual maturity at the onset of lay. The results from this thesis suggested that the genetic correlations between early, late and total production are positive, and thus the selection on early records may favour hens that are more mature by the time of the beginning of lay. However, decreasing the age at first egg can result in smaller eggs that are less likely to be hatchable. Nevertheless, the use of selection indices can assist breeders in balancing the benefits from using the early records with undesirable side effects.

The advantage that RR models offers is that they provide an effective method to account for the longitudinal fluctuations of genetic variance. Nevertheless, these changes could also reflect sampling error of the estimates. Sales and Hill (1976) have shown that the use of noisy information could inflate selection index. Nevertheless, the standard errors of the estimates of variance components from the monthly model were small and almost identical to the error from the multivariate model (see Table 5.2). Similarly, the standard errors of the genetic correlations from the bivariate model were close to the corresponding errors from Chapter 3 (see Table 3.5 and Table 7.5). The results from the daily model also suggest that the standard deviation of the posterior distribution of the genetic parameters was low (see Figure 6.3). Variance components from the random regression estimates were increased at both ends of the production period. This reflects the variable performance of birds in these stages, since at the onset of lay the initiation of egg production is affected by the sexual maturity and toward the end of the laying period, the persistency of lay determines the productivity. In these periods both genetic and phenotypic variance appeared increased when using a random regression model, but the application of a multi-variate model suggested the same.

In conclusion, RR models appear to offer some advantages for the genetic evaluation of egg production. The incorporation of a simple RR model, like the one developed in Chapter 5, in the routine evaluation of a breeding program is not anticipated to be difficult. Such models are used for routine genetic evaluation in dairy cattle for a number of years. Nevertheless, the extension to a daily threshold model may not be as easy, because of the model's increased requirements in computing resources. So, its implementation may not be possible for practical purposes for the time being. However, thanks to advances in computer technology, this constraint is likely to be alleviated in the foreseeable future.

Despite the proposed improvements for the genetic evaluation of egg production, the unfavourable genetic correlation between egg production and body weight remained an issue. Although this genetic association tended to have a dynamic character (Chapter 7), the evidence was not sufficient to pinpoint a laying stage presenting a breeding opportunity for both traits. Arguably, the magnitude of the correlation was prohibitive and therefore, it may be suggested that this extreme value was population-specific. So, in this challenging breeding situation, the best strategy to maintain the sustainability of selection is to maximise the efficiency of selection for both traits. In this thesis it was suggested that the application of longitudinal models can contribute towards this goal. Moreover, egg production currently appears to be of low priority compared to growth traits, so a selection index with appropriate weights can also provide a mechanism to meet the breeding objective.

In conclusion, the application of longitudinal genetic analysis of egg production may have considerable implications for poultry breeding. In the next section, a series of significant points that are expected to improve the output of selection for egg production are outlined.

8.3.3 Selecting for improved egg production

The key question that arises is what is the most appropriate selection method for improving egg production. In this thesis various suggestions have been made that are summarised below.

The most simple, yet necessary, step is the data collection on a longitudinal basis. In breeding companies eggs are collected on a daily basis, hence it is straight-forward to assign them to hens. The importance of digital recording (e.g. using barcoded tags) on farms is highlighted, since it minimises the occurrence of identification errors. Although the collection of data is performed at a considerable cost, the utilisation of the data into longitudinal models is still not applied on a practical basis. Therefore, there is scope for improvements in line with the conclusions from this thesis.

It is well known that multivariate analysis offers an array of advantages in terms of increasing the efficiency of selection (for a summary see section 7.1). For the total egg number the implementation is straightforward (see Chapter 3). Nevertheless, I showed it was feasible to perform a multivariate genetic analysis using RR for accounting for the dynamic character of laying. The benefit of the increased heritability estimates should translate to more rapid genetic progress for egg production. The same conclusion is expected to apply for a multivariate analysis employing a survival model, yet this cannot be easily implemented yet.

The RR framework allows targeted selection to the most informative laying stage, as discussed in section 8.3.2. In all analyses considered here, the economic value of eggs was assumed to be uniform throughout the laying period. However, with the use of different weights according to the timing of lay, it is possible to focus on interesting stages in terms of their economic importance. For instance, egg production decreases toward the ending of the laying period, but eggs are larger, yielding larger poults that might have an advantage in terms of survival and growth rates that translates to additional economic benefit for producers.

Breeding objectives are reviewed periodically and therefore a flexible analytical framework offers the machinery to adjust to new challenges and include additional traits. The importance of time patterns may not be restricted only to egg production. There are other traits that constitute components of the reproduction performance of a hen. Hence, for breeders it is not only important that a hen lays as many eggs as possible, but it is also essential that these eggs are fertile and hatchable. Particularly the improvement of late fertility/hatchability is a significant breeding objective. The implementation of longitudinal models for these traits can contribute in an analogous way as described in this thesis. In conclusion, as long as the longitudinal dimension of a trait is integrated in the genetic evaluation system, it can easily be extended to meet future challenges.

8.4 Future developments

This thesis drew attention to the longitudinal aspects of egg production and demonstrated their importance. However, there is a lack of similar studies in chicken, despite the potential to expand this methodology framework in this sector. It would be intriguing to repeat the analyses described here to an array of populations with different genetic backgrounds, such as laying hens and broiler breeders. This would allow comparisons between lines and/or species and thus it would improve our understanding of the relationship of laying patterns with other significant production traits.

This study implied that genetic factors had a time-depended effects. Thanks to the recent publication of the chicken genome (Hillier *et al.*, 2004), the availability of markers has been significantly increased, while the cost of genotyping dropped considerably. Thus, a handful of studies has been performed for the identification of QTLs (Hocking, 2005; Abasht *et al.*, 2006). Nevertheless, only the study of Minvielle *et al.* (2006) considered the longitudinal character of egg production. However, such studies would benefit from the application of a RR model, since it provides a robust framework to perform longitudinal studies. Therefore, it would be possible to detect QTLs with laying stage specific effects, as this thesis postulated.

In an animal breeding context, the implementation of a genomic approach can substantially accelarate the genetic improvement of egg production. It is known that for traits that are sex-limited and recorded later in life the use of Marker-Assisted Selection (MAS) can improve the accuracy of selection by permiting the combination of parental information and genotypic data from the poults as an indicator of their potential (Dekkers, 2004). The developement of a panel in which, for the markers included, associations for both egg production and growth traits are available, may also facilitate selection for those antagonistic traits. Extending the idea of using genotyping data, the breeding values of individuals can be estimated based solely on a large number of markers spread across the genome (Meuwissen *et al.*, 2001). Simulation studies have demonstrated the potential of genomic selection in dairy cattle, where the cost saved from the progeny testing is larger than the cost of genotyping (Schaeffer, 2006). In poultry though, because there is no progeny testing involved in the evaluation of birds, the cost of massive genotyping on a routine basis remains very high (although it will eventually be reduced). Therefore, for poultry breeders the challenge is to use effectively the genomic information and improve the output of the selection programmes.

8.5 General conclusions

The objective of this thesis was to investigate the genetics of egg production on a longitudinal basis. That being so, the focus was placed on accounting for time trends. Novel time-to-event traits, such as counting the time required for a hen to lay a fixed number of eggs, were formulated, that implied that the dynamic character of egg production is important. On a more detailed analysis it was demonstrated that the genetic variance affecting egg laying changes over time. The results suggested that the longitudinal character of variance components can be precisely described. Therefore, the treatment of egg laying as a longitudinal trait can provide a more accurate modelling of the underlying biological mechanisms. On a practical basis this can translate to increased efficiency of selection, due to the higher accuracy of prediction of breeding values. Furthermore, the detailed covariance structure obtained on a longitudinal basis can be used to target more effectively the selection pressure on the most informative stages of laying and thus to maximise the genetic gain from breeding programmes. However, it was also noted that complex models need to be applied for fully deploying the potential of this approach. Despite the complications, the adoption of a longitudinal framework for the genetic evaluation of reproduction traits in poultry appears to be favourable for the improvement of current breeding strategies.

Bibliography

Abasht, B., Dekkers, J. and Lamont, S.: 2006, Review of quantitative trait loci identified in the chicken, *Poultry Science* **85**, 2079:2096.

Ali, T. and Schaeffer, L.: 1987, Accounting for covariances among test day milk yields in dairy cows, *Canadian Journal of Animal Science* **67**, 637:644.

Anang, A., Millenz, N. and Schuler, L.: 2000, Genetic and phenotypic paramenters for monthly egg production in white leghorn hens, *Journal of Animal Breeding Genetics* **117**, 407:415.

Anang, A., Millenz, N. and Schuler, L.: 2001, Monthly modle for genetic evaluation of laying hens 1. fixed regression, *British Poultry Science* **42**, 191:196.

Anang, A., Millenz, N. and Schuler, L.: 2002, Monthly modle for genetic evaluation of laying hens ii. random regression, *British Poultry Science* **43**, 384:390.

Anthony, N., Dunnington, E. and Siegel, P.: 1989, duction and egg composition of parental lines and f and f1 2 duction and egg composition of parental lines and f and f1 2 egg production and egg composition of parental lines and f1 and f2 crosses of white rock chickens selected for 56-day body weight, *Poultry Science* **68**, 27:36.

Arthur, J. and Abplanalp, H.: 1975, Linear estimations of heritabilty and genetic corrrelation for egg production, body weight, conformation and egg weight of turkeys, *Poultry Science* **54**, 11:23.

Arthur, J. and Albers, G.: 2003, *Industrial perspective on problems and issues associated with poultry breeding*, CABI publishing, chapter 1, p. 1:12.

Averill, T., Rekaya, R. and Weigel, K.: 2006, Random regression models for male and female fertility evaluation using longitudinal binary data, *Journal of Dairy Science* **89**, 3681:3689.

Ayyagari, V., Mohaptra, S., Venkatramaiah, A., Thiagasundaram, T., Choudhuri, D., Johri, D. and Renganathan, P.: 1980, Selection for egg production on part records - part 1: evolution of short term response to selection, *Theoretical and Applied Genetics* **57**, 277:283.

Bayyari, G., Huff, W., Rath, N., Balog, J., Newberry, L., Villiness, J. D., Skeeles, J., Anthony, N. and Nestor, K.: 1997, Effect of the genetic selection of turkeys for increased body weight and egg production on immune and physiological responses, *Poultry Science* **76**, 289:296.

Berri, C., Wacrenier, N., Millet, N. and LeBihan-Duval, E.: 2001, Effect of selection for improved body composition on muscle and meat characteristics of broilers from experimental and commercial lines, *Poultry Science* **80**, 833:838.

Besbes, B. and Ducrocq, V.: 2003, *Use of mixed model methodology in breeding strategies for layers*, CABI publishing, chapter 9, p. 127:146.

Besbes, B., Ducrocq, V., Foulley, J., Protais, M., Tavemier, A., Tixier-Boichard, M. and Beaumont, C.: 1993, Box-cox transformation of egg-production traits of laying hens to improve genetic parameter estimation and breeding evaluation, *Livestock Production Science* **33**, 313:326.

Box, G. and Cox, D.: 1964, An analysis of transformations, *Journal of the Royal Statistical Society - Series B Statistical Methodology* **26**, 211:252.

Bradburn, M., Clark, T., Love, S. and Altman, D.: 2003, Survival analysis part ii: Multivariate data analysis – an introduction to concepts and methods, *British Journal of Cancer* **89**, 431:436.

Buchanan, S., Robertson, G. and Hocking, P.: 2000a, Developement of the reproductive system in turkeys with a high or low susceptibility to prolapse of the oviduct, *Poultry Science* **79**, 1491:1498.

Buchanan, S., Robertson, G. and Hocking, P.: 2000b, Effects of food restriction or delayed photostimulation on ovarian follicle number, plasmsa oestradiol concetration and vaginal collagen content in male-line turkeys, *British Poultry Science* **41**, 502:507.

Buchanan, S., Robertson, G. and Hocking, P.: 2002, Comparative changes in plasma concentrations of progesterone, oestradiol and lh during the ovulatory cycle in a multiple ovulating male line and a single ovulating traditional line of turkeys, *Reproduction* **123**, 127:133.

Buss, E.: 1990, *Genetics of growth and meat production in turkeys*, Elsevier, chapter 26, p. 645:677.

Chambers, J.: 1990, *Genetics of growth and meat production in chicken*, Elsevier, chapter 25, p. 599:643.

Chang, Y., D.Gianola, Heringstad, B. and Klemetsdal, G.: 2004, Longitudinal analysis of clinical mastitis at different stages of lactation in norwegian cattle, *Livestock Production Science* **88**, 251:261.

Chapuis, H., Tixier-Boichard, M., Delabrosse, Y. and V, D.: 1996, Multivariate restricted maximum likelihood estimation of genetic parameters for production traits in three selected turkey strains, *Genetics Selection Evolution* **28**, 299:317.

Clayton, G.: 1975, Normality of egg production in poultry, *British Poultry Science* **16**, 431:439.

Cox, D.: 1972, Regression models and life tables, *Journal of the Royal Statistical Society - Series B Statistical Methodology*.

Damgaard, L. and Korsgaard, I.: 2006, A bivariate quantitative genetic model for a linear gaussian trait and a survival trait, *Genetics Selection Evolution* p. 45:64.

Darwash, A., Lamming, G. and Woolliams, J.: 1997, The phenotypic association between the interval to post-partum ovulation traditional measures of fertility in dairy cattle, *Animal Science* **65**, 9:16.

Decuypere, E., Bruggeman, V., Onagbesan, O. and Safi, M.: 2002, Endocrine physiology of reproduction in the female chicken: Old wine in new bottles, *Avian and poultry biology reviews* **13**, 145–153.

Dekkers, J.: 2004, Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons, *Journal of Animal Science* **82**, 313:328.

Douglas, J. D.: n.d., Optimising performance of today's heavier breeder hen: "role of genetics", *Technical report*, Hybrid Turkeys.

Ducrocq, V.: 1987, *An analysis of length of productive life in dairy cattle*, PhD thesis, Cornell University, Ithaca, New York, USA.

Ducrocq, V. and Casella, G.: 1996, A bayesian analysis of mixed survival models, *Genetics Selection Evolution* **28**, 505:529.

Ducrocq, V. and Sölkner, J.: 1998, The survival kit—a fortran pack-age for the analysis of survival data, *Proceedings of 6th World Congress in Genetics Applied to Livestock Production*, p. 447:448.

Ducrocq, V., Besbes, B. and Protais, M.: 2000, Genetic improvement of laying hens viability using survival analysis, *Genetics Selection Evolution* **32**, 24:40.

Elhalawani, M. and Rosenboim, I.: 1993, The ontogeny and control of incubation behaviour in turkeys, *Poultry Science* **72**, 906:911.

Emmerson, D.: 2003, *Breeding objectives and selection strategies for broiler production*, CABI publishing, chapter 8, p. 113:126.

Emmerson, D., Velleman, S. and Nestor, K.: 2002, Genetics of growth and reproduction in the turkey: 15. effect of long-term of selection for increased egg production on the genetics of growth and egg production traits, *Poultry Science*.

Fairfull, R. and Gowe, R.: 1990, *Genetics of egg production in chickens*, Elsevier, chapter 29, p. 705:759.

Fairfull, R., McMillan, I. and Muir, W.: 1998, Poultry breeding: progress and prospects for genetic improvement of egg and meat production, *Proceedings of 6th World Congress in Genetics Applied to Livestock Production*, Vol. 24, p. 271:278.

Farhat, A., Lavallée, S., Chavez, E. R., Touchburn, S. P. and Lague, P. C.: 1998, Reproductive performance of f1 pekin duck breeders selected with ultrasound scanning for breast muscle thickness and the effect of selection on f2 growth and muscle measurement, *Technical report*, McGill University.

Fernandez, X., Sante, V., Baeza, E., LeBihan-Duval, E., Berri, C., Remignon, H., Babile, R., LePottier, G., Millet, N., Berge, P. and Astruc, T.: 2001, Post mortem muscle metabolism and meat quality in three genetic types of turkey, *British Poultry Science* **42**, 462:469.

Feurer, L.: 1957, The principle of simplicity, *Philosophy of Science* p. 109:122.

Fischer, T., Gilmour, A. and van der Werf, J.: 2004, Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information reml, *Genetics Selection Evolution* **36**, 363:369.

Flock, D.: 1977, Genetic analysis of part-period egg production in a population of white leghorns under rrs, *Z. Tierzucht Zuchtgsbiol* **94**, 89:103.

Flock, D.: 1998, Genetic-economic aspects of feed efficiency in laying hens, *World's Poultry Science* **54**, 225:239.

Gallo, L., Carnier, P., Cassandro, M., Zotto, R. D. and Bittante, G.: 2001, Test day genetic analysis of condition score and heart test day genetic analysis of condition score and heart girth in holstein friesian cows, *Journal of Dairy Science* **84**, 2321:2326.

Genstat: 2006, Genstat reference manual - version 9.

Gianola, D.: 2000, Statistics in animal breeding, *Journal of American Statistical Association* **95**, 296:299.

Gilmour, A. R., Cullis, B., Welham, S. and Thompson, R.: 2001, Asreml reference manual.

Groen, A.: 2003, *Breeding objectives and selection strategies for layer production*, CABI publishing, chapter 7, p. 101:112.

Grossman, M. and Koops, W.: 2001, A model for individual egg production in chickens, *Poultry Science* **80**, 859:867.

Grossman, M., Grossman, T. and Koops, W.: 2000, A model for persistency of egg production, *Poultry Science* **79**, 1715:1724.

Havenstein, G.: 2003, Progress in broiler breeding: How far can we go?, *3rd European poultry genetics symposium*, p. 10:17.

Havenstein, G., Ferket, P. and Qureshi, M.: 2003, Growth, livability, and feed conversion of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets, *Poultry Science* **82**, 1500:1508.

Heringstad, B., Rekaya, R., Gianola, D., Klemetsdal, G. and Weigel, K.: 2003, Genetic change for clinical mastitis in norwegian cattle: a threshold model, *Journal of Dairy Science* **86**, 369:375.

Hillier, L., Miller, W. and Birney, E.: 2004, Sequence and comparative analysis of the chicken genome provide unique perspective on vertebrate evolution, *Nature* **432**, 695:716.

Hocking, P.: 1993, Relationship between egg size, weight and pelvic dimensions in turkeys, *Animal production* **56**, 145:150.

Hocking, P.: 2003, Comparative effects of feed restriction in broiler and turkey breeders, *Proceedings of 5th International Symposium on Turkey Reproduction*, North Carolina University, p. 36:41.

Hocking, P.: 2005, Review of qtl mapping results in chickens, *World's Poultry Science* **61**, 215:226.

Hocking, P. and Bernard, R.: 1998, Comparative developement of the ovary and production, fertility and hatchability of eggs from traditional turkeys and a contemporary male-line fed ad libitum or restricted, *British Poultry Science* **39**, 291:297.

Hocking, P., Gilbert, A. B., Whitehead, C. and Walker, M.: 1988, Effects of age and of early lighting on ovarian function in breeding turkeys, *British Poultry Science* **29**, 223:229.

Hocking, P., Mitchell, M. A., Bernard, R. and Sandercock, D. A.: 1998, Interaction of age, strain, sex and food restriction on plasma creatine kinase activity in turkeys, *British Poultry Science* **39**, 360:364.

Hocking, P., Waddington, D. and Walker, M.: 1992, Changes in ovarian function of female turkeys photostimulated at 18, 24 or 30 weeks of age and fed ad libitum or restricted until point of lay, *British Poultry Science* **33**, 639:647.

Hofer, A.: 1998, Variance component estimation in animal breeding: a review, *Journal of Animal Breeding Genetics* **115**, 247:265.

Hort, J., Knizetova, H. and Hyankova, L.: 2000, Selection for high body weight and correlated responses in reproduction in poultry, *Czech journal of animal science* **45**, 519:523.

Huisman, A., Kanis, E. and van Arendonk, J.: 2002, Application of random regression models in pig breeding to select on growth and feed intake patterns, *Proceedings of 53th meeting of European Association for Animal Production*, p. 1:8.

Ibe, S. and Hill, W.: 1988, Transformation of poultry egg production data to improve normality, homoscedasticity and linearity of genotypic regression, *Journal of Animal Breeding Genetics* **105**, 231:240.

Jaffrezic, F., Thompson, R. and Pletcher, S.: 2004, Multivariate character process models for the analysis of two or more correlated function-valued traits, *Genetics* **168**, 477:487.

Jensen, J. and Madsen, P.: 2006, Dmu: A package for the analysis of multivariate mixed models, *Proceedings of 8th World Congress in Genetics Applied to Livestock Production*.

Jones, K.: 2006, Animal products markets in 2005 and forecasts for 2006 animal products markets in 2005 and forecasts for 2006 animal products markets in 2005 and forecasts for 2006 animal products markets in 2005 and forecasts for 2006 animal product markets and in 2005 and forecast for 2006, *Technical report*, United States Department of Agriculture.

Kaplan, E. and Meier, P.: 1958, Onparametric estimation from incomplete observations, *Journal of American Statistical Association* **53**, 457:481.

Karacaören, B., Jaffrézic, F. and Kadarmideen, H. N.: 2006, Genetic parameters for functional traits in dairy cattle from daily random regression models, *Journal of Dairy Science* **89**, 791:798.

Katanbaf, M., Dunnington, E. and Siegel, P.: 1989, Restricted feeding in early and late-feathering chickens. 2. reproductive responses, *Poultry Science* **68**, 352:358.

Kerr, C. L., Hammerstedt, R. H. and Barbato, G. F.: 2001, Effects of selection for exponential growth rate at different ages on reproduction in chickens, *Avian and poultry biology reviews* **12**, 127:134.

Kinghorn, B. and Kinghorn, S.: 2006, Pedigree viewer program (version 5.4).

Kirkpatrick, M. and Heckman, N.: 1989, A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters, *Journal of Mathematical Biology* **27**, 429:450.

Kirkpatrick, M., Hill, W. and Thompson, R.: 1994, Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle, *Genetical Research* **64**, 57:69.

Kirkpatrick, M., Lofsvold, D. and Bulmer, M.: 1990, Analysis of inheritance, selection and evolution of growth trajectories, *Genetics* **124**, 221:240.

Klein-Hessling, H.: 2003, Body weight control of turkey breeder hens - an international perspective, *Proceedings of 5th International Symposium on Turkey Reproduction*, North Carolina University, p. 18:23.

Kleinbaum, D. and Klein, M.: 2002, *Logistic regression - A self learning text*, second edition edn, Springer-Verlang, New York.

Kleinbaum, D. and Klein, M.: 2005, *Survival analysis - A self learning text*, Springer-Verlang.

Koerhuis, A.: 1996, Non-normality of egg production distributions in poultry the effects of outlier elimination transformation on size curvilinearity of heritability, *Livestock Production Science* **45**, 69:85.

Koerhuis, A., Mckay, J., Hill, W. and Thompson, R.: 1997, A genetic analysis of egg quality traits their maternal influence on offspring-parental regressions of juvenile body weight performance in broiler chickens, *Livestock Production Science* **49**, 203:215.

Korsgaard, I., Andersen, A. and Jensen, J.: 1998, Bayesian inference in the semiparametric log normal frailty model using gibbs sampling, *Genetics Selection Evolution* **30**, 241:256.

LeBihan-Duval, E., Berri, C., Baeza, E., Millet, N. and Beaumont, C.: 2001, Estimation of the genetic parameters of meat characteristics and of their genetic correlations with growth and body composition in an experimental broiler line, *Poultry Science* **80**, 839:843.

LeBihan-Duval, E., Mignon-Grasteau, S., Millet, N. and Beaumont, C.: 1998, Genetic analysis of a selection experiment on increased body weight and breast muscle weight as well as on limited abdominal fat weight, *British Poultry Science* **39**(3), 346:353.

LeBihan-Duval, E., Millet, N. and Remignon, H.: 1999, Broiler meat quality: effect of selection for increased carcass quality and estimates of genetic parameters, *Poultry Science* **78**, 822:826.

Li, Z., Nestor, K., Saif, Y., Andreson, J. and Patterson, R.: 2001, Effect of selection for increased body weight in turkeys on lymphoid organ weights, phagocytosis, and antibody responses to fowl cholera and newcastle disease-inactivated vaccines, *Poultry Science* **81**, 689:694.

Lynch, M. and Walsh, B.: 1998, *Genetics and analysis of quantitative traits*, Sinauer.

Marks, H.: 1990, *Genetics of egg production in other galliforms*, Elsevier, chapter 30, p. 761:770.

McCartney, M., Nestor, K. and Harvey, W.: 1968, Genetics of growth and reproduction in the turkey: 2. selection for increased body weight and egg production, *Poultry Science* **47**, 981:990.

McMillan, I., Gowe, R., Gavora, J. and Fairfull, R.: 1986, Prediction of annual production from part record egg production in chickens by three mathematical models, *Poultry Science* **65**, 817:822.

Melnychuk, V., Robinson, F., Renema, R. and Hardin, R.: 1997, Carcass traits and reproductive developement at the onset of laying in two lines of female turkeys, *Poultry Science* **76**, 1197:1204.

Meuwissen, T., Hayes, B. and Goddard, M.: 2001, Prediction of total genetic value using genome-wide dense marker maps, *Genetics* **157**, 1819:1829.

Meyer, K.: 1989, Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm, *Genetics Selection Evolution* **21**, 317:340.

Meyer, K.: 1998, Estimating covariance functions for longitudinal data using a random regression model, *Genetics Selection Evolution* p. 221:240.

Meyer, K. and Hill, W.: 1997, Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood, *Livestock Production Science* p. 185:200.

Minvielle, F., Kayang, B., Inoue-Murayama, M., Miwa, M., Vignal, A., Gourichon, D., Neau, A., Monvoisin, J.-L. and ichi Ito, S.: 2006, Search for qtl affecting the shape of the egg laying curve of the japanese quail, *BMC Genetics*.

Mrode, R.: 2005, *Linear models for the prediction of animal breeding values*, 2nd edition edn, CABI publishing.

Nestor, K.: 1976, Selection for increased semen yield in the turkey., *Poultry Science* **55**, 2363:2369.

Nestor, K.: 1977a, Genetics of growth and reproduction in the turkey: 5. selection for increased body weight alone and in combination with increased egg production, *Poultry Science* **56**, 337:347.

Nestor, K.: 1977b, The influence of a genetic change in egg production, body weight, fertility or response to cold stress on semen yield in the turkey, *Poultry Science* **56**, 421:425.

Nestor, K.: 1985, Genetics of growth and reproduction in the turkey: 10. tandem selection for increased body weight alone and egg production, *Poultry Science* **64**, 2221:2222.

Nestor, K., Andreson, J. and Patterson, R.: 2000, Genetics of growth and reproduction in the turkey: 14. changes in genetic parameters over thirty generations of selection for increased body weight, *Poultry Science* **79**, 445:452.

Nestor, K., Andreson, J. and Velleman, S.: 2001, Genetic variation in pure lines and crosses of large-bodied turkey lines. 2. carcass traits and body shape, *Poultry Science* **80**, 1093:1104.

Nestor, K., Noble, D., J.Zhu and Moritsu, Y.: 1996, Direct and correlated responses to long-term selection for increased body weight and egg production, *Poultry Science* **75**, 1180:1191.

Olori, V., Hill, W., McGuirk, B. and Brotherstone, S.: 1999, Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model, *Livestock Production Science* **61**, 53:63.

Otsu, K., Khanna, V., Archibald, A. and MacLennan, D.: 1991, Cosegregation of porcine malignant hyperhermia and a probable causal mutation in the skeletal muscle ryanodine receptor gene in backcross families, *Genomics* **11**, 744:750.

Ottinger, M.: 1992, Altered neuroendocrine mechanisms during reproductive ageing, *Poultry Science Reviews* **4**, 235:248.

Owens, C., Hirschler, E., McKee, S., Martinez-Dawson, R. and Sams, A.: 2000, The characterization and incidence of pale, soft, exudative turkey meat in a commercial plant, *Poultry Science* **79**, 553:558.

Peto, R. and Peto, J.: 1972, Asymptotically efficient rank invariant procedures, *Journal of the Royal Statistical Society* **A135**, 185:207.

Piles, M., Garreau, H., Rafela, O., Larzul, C., Ramon, J. and Ducrocq, V.: 2006, Survival analysis in two lines of rabbits selected for reproductive traits, *Journal of Animal Science* **84**, 1658:65.

Pool, M., Janss, L. and Meuwissen, T.: 2000, Genetic parameters of legendre polynomials for first parity lactation curves, *Journal of Dairy Science* **83**, 2640:2649.

Ptak, E. and Schaeffer, L.: 1993, Use of the test day yields for genetic evaluation of dairy sires and cows, *Livestock Production Science* **34**, 23:34.

R: 2006, The r project for statistical computing - version 2.4.0.

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

Reddy, R.: 1996, The effects of long-term selection on growth of poultry, *Poultry Science* **75**, 1164:1167.

Rekaya, R., Gianola, D., Weigel, K. and Shook, G.: 2003, Longitudinal random effects models for genetic analysis of binary data with application to mastitis in dairy cattle, *Genetics Selection Evolution* **35**, 457:468.

Rekaya, R., Rodriguez-Zas, S., Gianola, D. and Shook, G.: 1998, Test-day models for longitudinal binary responses: an application to mastitis in holsteins, *in* W. Press (ed.), *Proceedings of 49th meeting of European Association for Animal Production*, p. 44.

Remignon, H. and LeBihan-Duval, E.: 2003, *Meat quality problems associated with selection for increased production*, CABI publishing, chapter 4, p. 53:66.

Renema, R. and Robinson, F.: 2004, Defining normal: comparison of feed restriction and full feeding of female broiler breeders, *World's Poultry Science* **60**, 508:522.

Robertson, A. and Lerner, I.: 1949, The heritability of all-or-none traits: viability of poultry, *Genetics* **34**, 395:411.

Saebø, S. and Frigessi, A.: 2004, A genetic and spatial bayesian analysis of mastitis resistance, *Genetics Selection Evolution* **36**, 527:542.

Sales, J. and Hill, W.: 1976, Effect of sampling errors on efficiency. of selection indices. 2. use of information on associated traits, *Animal production* **23**, 1:14.

Sapp, R. L., Rekaya, R., Misztal, I. and Wing, T.: 2004, Male and female fertility and hatchability in chickens: A longitudinal mixed model approach, *Poultry Science*.

Savas, T., Preisinger, R., Rohe, R. and Kalm, E.: 1999, The effect of the box-cox-transformation on the estimation of genetic parameters for egg production traits in laying hens, *Zuchtungskunde* **71**, 380:391.

Schaeffer, L.: 2004, Application of random regression models in animal breeding, *Livestock Production Science*.

Schaeffer, L.: 2006, Strategy for applying genome-wide selection in dairy cattle, *Journal of Animal Breeding Genetics* **123**, 218:223.

Schneider, M.: 2006, *Fertility, Mastitis and Longevity in Dairy Cattle Analyzed Using Survival Models*, PhD thesis, Swedish University of Agricultural Sciences, Uppsala.

Schneider, M., Strandberg, E., Ducrocq, V. and Roth, A.: 2005, Survival analysis applied to genetic evaluation for female fertility in dairy cattle, *Journal of Dairy Science* **88**, 2253:2259.

Sölkner, J. and Ducrocq, V.: 1999, The survival kit: a tool for analysis of survival data, *in* I. Bulletin (ed.), *Workshop on Genetic Improvement of Functional Traits in Cattle - Longevity, Jouy-en-Josas, France*, Vol. 21, p. 181:189.

Sorensen, D. A. and Gianola, D.: 2002, *Likelihood, Bayesian, and McMc Methods in Quantitative Genetics.*, Springer-Verlang.

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

Southey, B., Rodriguez-Zas, S. and Leymaster, K.: 2000, Survival analysis of lamb mortality in a terminal sire composite population, *Journal of Animal Science* **79**, 2298:2306.

SPSS: 2006, Spss reference manual - version 14.

StataCorp: 2006, *STATA Reference manual - Version 9*, StataCorp LP.

Swatland, H.: 1989, Morphometry of pectoral developement in turkey breeding stock, *British Poultry Science* **30**, 785:795.

Szwaczkowski, T.: 2003, *Use of mixed model methodology in poultry breeding: estimation of genetic parameters*, CABI publishing, chapter 11, p. 165:202.

Tablante, N. L., Vaillancourt, J., Martin, S. W., Shoukri, M. and Estevez, I.: 2000, Spatial distribution of cannibalism mortalities in commercial laying hens, *Poultry Science*.

Taha, F.: 2003, The poultry sector in middle-income countries and its feed requirements: The case of egypt, *Technical report*, United States Department of Agriculture.

Thompson, R. and Meyer, K.: 1986, A review of theoretical aspects in the estimation of breeding values for multi-trait selection, *Livestock Production Science* **15**, 299:313.

Togashi, K. and Lin, C.: 2006, Selection for milk production and persistency using eigenvectors of the random regression coefficient matrix, *Journal of Dairy Science* **89**, 4866:4873.

Unver, Y., Akbas, Y. and Ogus, I.: 2004, Effect of box-cox transformation on genetic parameter estimation in turkeys, *Turkish Journal of Veterinary Animal Science* **28**, 249:255.

USDA: 2006, Poultry: 2006 forecast overview.

van Arendonk, J. and Bijma, P.: 2003, Factors affecting commercial application of embryo technologies in dairy cattle in europe - a modelling approach, *Theriogenology* p. 635:649.

van Middelkoop, J. and Siegel, P.: 1976, Classification of abnormal chicken egg, *Poultry Science* **55**, 1563:1566.

Vaupel, J., Manton, K. and Stallard, E.: 1979, The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* **16**, 439:454.

Veerkamp, R., Koenen, E. and DeJong, G.: 2001, Genetic correlations among body condition score, yield, and fertility in first parity cows estimated by random regression models, *Journal of Dairy Science* **84**, 2327:2335.

Velleman, S. and Nestor, K.: 2004, Inheritance of breast muscle morphology in turkeys at sixteen weeks of age, *Poultry Science* **83**, 1060:1066.

Vukasinovic, N., Moll, J. and Kunzi, N.: 1999, Genetic evaluation for the length of productive life with censored records, *Journal of Dairy Science* **82**, 2178:2185.

Wang, T., Fernando, R. and Kachman, S.: n.d., Matvec users' guide, version 1.03.

Werf, H. V. D., Goddard, M. and Meyer, K.: 1998, The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records, *Journal of Dairy Science* **81**, 3300:3308.

Whitehead, C., Fleming, R. H., Julian, R. and Sorensen, P.: 2003, *Skeletal problems associated with selection for increased production*, CABI publishing, chapter 3, p. 29:52.

Whitson, D., Mardsen, S. and Titus, H.: 1944, A comparison of the performance of four varieties of turkeys during the breeding season, *Poultry Science* **23**, 314:320.

Wilkiewicz-Wawro, E., Wawro, K., Lewczuk, A. and Milchalik, D.: 2003, Correlation between the thickness of breast muscles and meatiness in turkeys, *Czech journal of animal science* **48**, 216:222.

Wilson, B. W., Nieberg, P. and Buhr, R.: 1990, Turkey muscle growth and focal myopathy, *Poultry Science* **69**, 1553:1562.

Woolliams, C., Suttle, N., Woolliams, J., Jones, D. and Wiener, G.: 1986, Studies on lambs from lines genetically selected for low and high copper status. 1. differences in mortality, *Animal production* **43**, 293:301.

Wu, J., Zelinski, M., Ingram, D. and Ottinger, M.: 2005, Ovarian aging and menopause: current theories, hypotheses, and research models., *Experimental Biology and Medicine (Maywood N.J)* **230**, 818:828.

Yazdi, M., Rydhmer, L., Ringmar-Cederberg, E., Lundeheim, N. and Johansson, K.: 2000, Genetic study of longevity in swedish landrace sows, *Livestock Production Science* **63**, 255:264.

Yazdi, M., Visscher, P., Ducrocq, V. and Thompson, R.: 2002, Heritability, reliability of genetic evaluations and response to selection in proportional hazard models, *Journal of Dairy Science* **85**, 1563:1577.