

Exploring the Use of Random Regression Models with Legendre Polynomials to Analyze Clutch Size in Iranian Native Fowl

Research Article

A. Ebadi Tabrizi¹, M. Tahmoorespur^{1*} and A. Nejati Javaremi²

¹ Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

² Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

Received on: 7 Jun 2014

Revised on: 4 Aug 2014

Accepted on: 14 Aug 2014

Online Published on: Jun 2015

*Correspondence E-mail: tahmoores@um.ac.ir

© 2010 Copyright by Islamic Azad University, Rasht Branch, Rasht, Iran

Online version is available on: www.ijas.ir

ABSTRACT

Random regression models (RRM) have become common for the analysis of longitudinal data or repeated records on individual over time. The goal of this paper was to explore the use of random regression models with orthogonal / Legendre polynomials (RRL) to analyze new repeated measures called clutch size (CS) as a meristic trait for Iranian native fowl. Legendre polynomial functions of increasing order 0 (no covariate) to 4 were fitted to the age at sexual maturity (ASM) and 1 to 10 to the additive genetic and permanent environmental effects. Days in production (clutch) were used as time variables. Homogeneity of residual variance through the time was assumed. Analyses were carried out within restricted maximum likelihood algorithm (REML) using WOMBAT software. Adequacy of models was checked by Bayesian information criterion (BIC). The resulted BICs suggested a model composed of the second order polynomial for ASM and 8th order polynomial for additive genetic and permanent environmental effect was the most suitable for adjusting the present records. The highest phenotypic and permanent environmental variance of CS was at the beginning of the production period. Additive genetic variance was fairly consistent during 210 and 265 days of age (d 210-d265). Estimates of heritability for CS ranged from 0.033 to 0.199 for d 161 and d 242 in the first cycle of egg production, respectively. The ratio of animal permanent environmental variance to phenotypic variance was in the range of 0.01 and 0.264. The estimated ranges for additive genetic and permanent environmental correlations were -0.18 to 0.99 and -0.5 to 0.99, respectively and were high between the adjacent ages and they tended to decrease at nonadjacent ages.

KEY WORDS clutch size, Iranian native fowl, Legendre polynomial, meristic trait, random regression.

INTRODUCTION

Egg production in laying hens is a process of pronounced cyclic nature and eggs are laid at intervals of around 24 to 27 hours or more, depending on age (Akbas *et al.* 2002). These cyclic processes result in the formation of the clutches of one or more eggs laid on consecutive days and each clutch is followed by one or more pause days. The length of a clutch varies little within a hen (Koops *et al.* 1992). Clutch size, like litter size, is a type of quantitative trait called meristic (Rowland *et al.* 1997) and quantitative trait

methodologies can be applied to these kinds of traits (Crow *et al.* 2000).

Measurements that are taken along a trajectory of time can be modeled as a function of that trajectory by means of random regression models. The prime advantage of studying longitudinal records like CS using RRM is its effectiveness for studying change (Van der Werf, 2001).

Hens with the same total or partial production, however, can exhibit different egg production curves (patterns) because of differences in persistency (Grossman *et al.* 2000). The effectiveness of selection can be improved by deeper

understanding and more detailed analysis of the laying process (Koops *et al.* 1992). Therefore, the idea to use part-time records as separate or repeated measurement of egg production as proposed by Wolc *et al.* (2007), Wolc *et al.* (2009), Wolc *et al.* (2010), Wolc *et al.* (2011), Wolc *et al.* (2013), Farzin *et al.* (2010), Kranis *et al.* (2007), Anang *et al.* (2001) and Anang *et al.* (2002) leads to loss of pattern and characterization of egg production. The use of RRM for the genetic evaluation of clutch size has not been reported and studies focusing especially on the order of polynomials can be performed using these models. The main objective of this study was to find the best random regression model and estimate the (co) variance components and variance ratios including heritability and ratio of animal permanent environmental variance to phenotypic variance of the clutch size as a new repeated record. To find the relative trajectory of variance components of CS according to the mean egg production of the flock, the hen-day curve was used.

MATERIALS AND METHODS

Egg production records were collected from 3156 birds comprising 10 generations of pedigreed birds from Iranian native fowl production center in Mazandaran province. Number of sires and dams were 81 and 516, respectively. After hatching, chicks were transferred to rearing barns.

Every 10 hens were kept in one pen and hens laid in individual trap nests, and production was individually recorded on daily basis for 365 days. The records are consisted of a sequence of 365 binary records (0=no egg, 1=egg laid), with each record corresponding to a specific day of the whole production period. If a hen died during the laying period, the rest of its records after the date of the death were treated as missing values. Cracked eggs were counted as laid eggs.

Eggs laid on the floor were excluded, because they could not be assigned to a specific hen. The consecutive days of egg production without any delay is called CS. The first day of each clutch was assigned as test-day of that clutch. Data consists of 111,137 test-days CS which was extracted from daily egg production.

The trait analyzed in the current study was the CS which was extracted from eggs laid in trap nests over the period of 365 days. In order to edit and explore fixed effects on the trait R was used. The data were analyzed using the random regression model. All the analyses were performed using PXEM restricted maximum likelihood algorithm with the WOMBAT software (Meyer, 2006). Fifty alternative RRM were analyzed.

All models include in their equation the same systematic effects; and temporary measurement errors were taken to be

independently distributed with constant variance along days in production (DIP).

The analysis of CS was based on the following RRM:

$$Y_{ijk} = Hatch_i + \sum_{j=1}^{n_1} \beta_j X_{jk} + \sum_{j=1}^{n_2} \alpha_j X_{jk} + \sum_{j=1}^{n_3} \psi_j X_{jk} + \epsilon_{ijk}$$

Where:

Y_{ijk} : K^{th} test-day observation of the j^{th} hen.

$Hatch_i$: independent fixed effect of i^{th} hatch.

β_j : j^{th} fixed regression coefficient.

α_j : j^{th} random regression coefficient for additive genetic effect.

ψ_j : j^{th} random regression coefficient for permanent environmental effect.

X_{jk} : covariate which is the value of Legendre polynomial on DIP.

ϵ_{ijk} : random residual effect.

n_1, n_2, n_3 : numbers of covariates, dependent on the order of Legendre polynomials.

The following structure of (co) variance matrices for random effects was assumed:

$$V \begin{pmatrix} \alpha \\ \psi \\ \epsilon \end{pmatrix} = \begin{pmatrix} G \otimes A & 0 & 0 \\ 0 & P \otimes I & 0 \\ 0 & 0 & E \otimes I \end{pmatrix}$$

Where:

G: genetic covariance matrix among random regression coefficients and the trait.

A: additive numerator relationship matrix.

P: permanent environmental covariance matrix among random regression coefficients and the trait.

E: residual variance assumed to be constant throughout the period for simplification.

\otimes : kronecker product.

Following Kirkpatrick *et al.* (1990), regression coefficients were derived from Legendre polynomials. The use of Legendre polynomials for RRM in monthly egg production was also recommended by Wolc *et al.* (2009). (Co) variance components for all models were estimated with the restricted maximum likelihood, employing PXEM algorithm, using WOMBAT program. Genetic and permanent environmental correlations were predicted for five values of the control variable, corresponding to lowest and highest value (d 116, d 407 respectively) and 3 approximately equidistant intermediate values (d 188, d 261, d 334; Meyer, 2006). Following Strabel *et al.* (2005), to compare the models, Bayesian Information Criteria (BIC), penalized-likelihood criteria, is used. Subsequently the lowest value of BIC was considered as best model (Venturini *et al.*

2012). The starting point of each curve was defined to be 116 which is the age at 5% egg production of the flock. The hen-day rates of the flock were calculated for each day.

RESULTS AND DISCUSSION

Estimated BIC for 50 models fitted for the present records is shown in Figure 1. According to estimated BICs a model composed of the second order polynomial for ASM and 8th order polynomial for additive genetic and permanent environmental effect was found to be the most suitable for adjusting the present records. The results were similar to studies carried out by Pool *et al.* (1999) and Wolc *et al.* (2009) who reported that model adequacy was better for models with higher order polynomials.

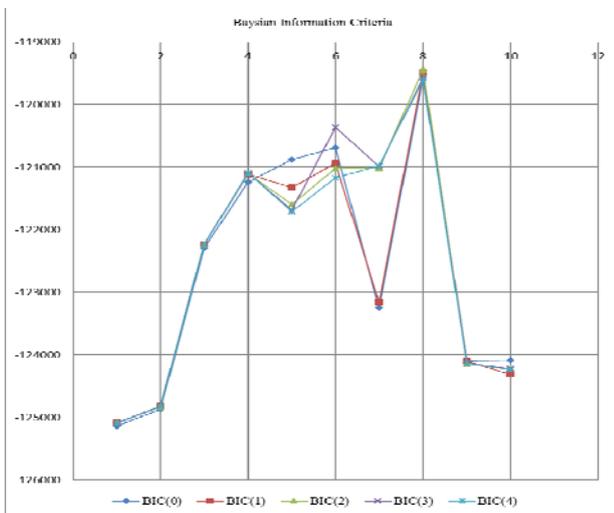


Figure 1 Bayesian information criteria (Y-axis) against order of fit for random regression effects (additive and permanent environmental effects) in X-axis, to check the adequacy of the models. BIC(0), BIC(1), BIC(2), BIC(3), BIC(4) stands for 1st, 2nd, 3rd, 4th order of fit for fixed regression effect (ASM), respectively

The range of hen-day index was between 0.05 and 0.57 which is shown in Figure 2. The curve indicates that the production period is consisted of two consecutive egg production cycles which the first one is complete (d 116 to d 371) but the second one is about 30 days. The first cycle is consisted of two main parts including a fast rise from the age at 5% egg production of the flock to peak production (d 144) and a two-phase decrease from peak to the end of the cycle. This pattern of egg production agrees with the reports of Miyoshi *et al.* (1996) and is not a typical egg production curve. A typical egg production curve for a flock increases rapidly and then decrease at a constant rate to the end of the production period (Grossman, 2000) but in the current curve the decreasing part had a long slow decrease and then a short steep decrease. The duration of these three stages were 28, 210 and 17 days, respectively.

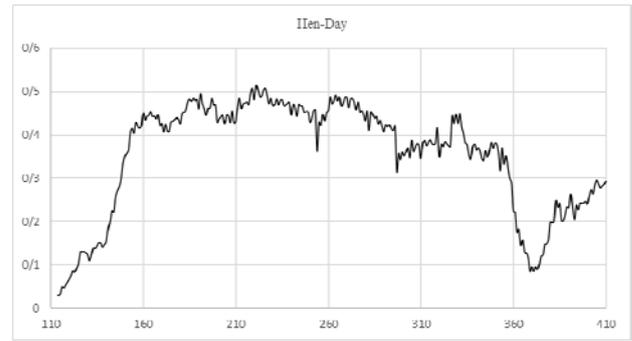


Figure 2 Hen-day against age of the flock in days

Phenotypic, permanent environmental and additive genetic variances of CS are shown in Figure 3. The highest phenotypic and permanent environmental variances of CS were at the first stage of the first cycle. The pattern of phenotypic variance at the starting point was similar to the results gained by Venturini *et al.* (2012) for monthly egg production but the highest additive genetic variance occurred in the second stage of the first cycle which is different from the results of mentioned study. The overall trend of phenotypic and permanent environmental variances was decreasing and the overall trend for additive genetic variance had two stages; 1) increasing from the beginning of the egg production period (d 116) to a peak (d 242), 2) decreasing from the peak to the end of the production period. Oscillatory nature of phenotypic, additive genetic and permanent environmental variances agreed with the results gained by Lukovic *et al.* (2003) on number of piglets born alive (NBA) which is another example of meristic trait. As mentioned by Lukovic *et al.* (2003) at the end of trajectory, decrease in phenotypic, additive genetic and permanent environmental variances probably were the consequence of small number of data. As proposed by Meyer (1998), some scaling of the observed ages can be considered, for instance, a logarithmic transformation should reduce the impact of few observations on very old individuals which is not applied to the current records. According to the changes of phenotypic, additive and permanent environmental variances, variations of phenotypic effects at the first stage of hen-day curve are because of permanent environmental effects. Heritability of CS was in the range of 0.033 to 0.199 (Figure 4) which was similar to the average estimated heritability obtained in the research conducted by Wolc *et al.* (2013) for pedigree model (0.2) and was somehow similar to the range of heritability estimated by Venturini *et al.* (2012) that was from 0.04 to 0.14. The lowest and the highest heritability estimates were at beginning and middle part of second phase of hen-day curve respectively (d 163, d 242) and this pattern of changes in heritability were different from results in Wolc *et al.* (2009) and Farzin *et al.* (2010) which highest heritability were at first month of egg

production. The pattern of changes in heritability were also different from the results of study conducted by [Kranis et al. \(2006\)](#) which the random regression using second-order Legendre polynomial was fitted to five monthly egg production records in Turkey. According to results obtained by [Kranis et al. \(2006\)](#) the highest and the lowest estimates of heritability occurred in the first and second monthly egg production records, respectively. Changes of heritability over time may result from activation of different genes during production cycle. Early stages of production are under the influence of sexual maturity ([Wolc et al. 2009](#)) but from the point of peak production, genes related to persistency of egg production could be more influential ([Minvielle et al. 2006](#)).

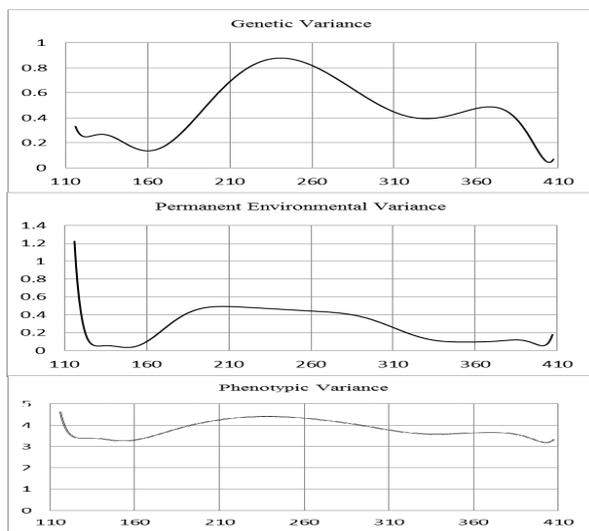


Figure 3 Changes of additive genetic, permanent environmental and phenotypic variance against age in days

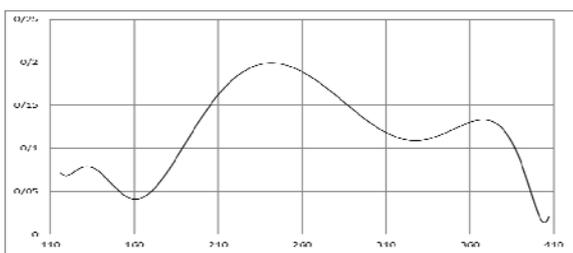


Figure 4 Estimated heritability from the most adequate model against age in days

The ratio of permanent environmental variance to phenotypic variance was in the range of 0.01 to 0.264 (Figure 5). The overall trend of the ratio was decreasing which agreed with the results gained by [Kranis et al. \(2007\)](#). At the beginning of the curve there was a rapid downward and upward change which is occurred around the peak production to d 185. Next, the ratio had a slow decrease to the end part of the first cycle. At the end of the curve which is simulta-

neous to the beginning of the second cycle of egg production the curve had smooth decreasing oscillations. As [Lukovic et al. \(2004\)](#) shown although the overall trend of the ratio was increasing but smooth oscillations have been observed.

The estimated genetic correlations for five values of the control variables, corresponding to lowest and highest values (d 116, d 407, respectively), and three approximately equidistant intermediate values (d 188, d 261, d 334) are presented in Figure 6. For the most adequate model, the range of genetic correlations for five values of the control variables including d 116, d 188, d 261, d 334 and d 407 were 0.258 to 0.99, 0.294 to 0.99, -0.018 to 0.99, -0.058 to 0.99 and -0.18 to 0.99 respectively. The correlations for d116 and d188 were all positive. The lowest correlations are found for d 407. The genetic correlations for three middle values including d 188, d 261 and d 344 were dominantly greater than 0.5. As the distance between the age of 116 and the consecutive ages increased the genetic correlations decreased rapidly and this pattern continued to the peak production of the first cycle and then the correlations had a slow increasing trend up to the beginning of the second cycle. The beginning of the first and second cycle of egg productions had high genetic correlations which is the indicator of common gene activations. The genetic correlations for the age of 188 decreased as the distant between consecutive days increased. The slope of the curve for the ages after d 188 was less than the ages before this point which means that changes in genetic correlations for the days after the point of d 188 were less than the days before this point. The correlations between this point and the beginning of the second cycle were high. Changes of the genetic correlations for the age of 261 as the middle point of the whole egg production period, had a slow decrease as the distance between consecutive points increased and these patterns were consistent in two opposite directions; 1) from this point to the peak production of the first cycle and 2) from this point to the beginning of the second cycle. The correlations between this point and the beginning of the first and second cycle decreased fast. The correlations for the ages of 334 and 407 as two end point of the curve had the same patterns as d 116 and d 188 but in the reverse direction.

The genetic correlations for partial egg production which are conducted by [Venturini et al. \(2012\)](#) and [Wolc et al. \(2007\)](#) were in the ranges of 0.1 to 0.99 and -0.177 to 0.99, respectively and the results of the current study were similar to [Wolc et al. \(2007\)](#) and different from [Venturini et al. \(2012\)](#). The overall pattern of changes of genetic correlations in the current study indicated that as the distance between adjacent ages increased the genetic correlations decreased and patterns agreed with the results obtained by

Komprej *et al.* (2013) for milk traits in Slovenian dairy sheep including milk yield, fat content and protein content, Lukovic *et al.* (2007) for litter size in pig, Wolc *et al.* (2007) for monthly egg production, Akbas *et al.* (2004) for quail body weights and Jensen (2001) for milk yield in dairy cattle. Despite the fact that at the time of this study, no research has been done on the architecture of CS as a time-dependent trait, the patterns of genetic correlations for five values of the control variables indicated probably the same genetic architecture for beginning point of the two consecutive egg production cycles and different genetic architecture for the point of peak production and the beginning of the two consecutive cycles. Estimates of permanent environmental correlations are also given in figure 6.

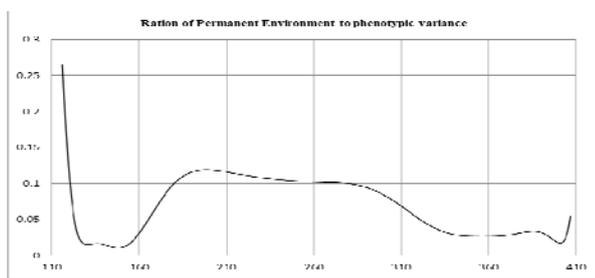


Figure 5 Estimated ratio of animal permanent environment to phenotypic variance from the most adequate model against age in days

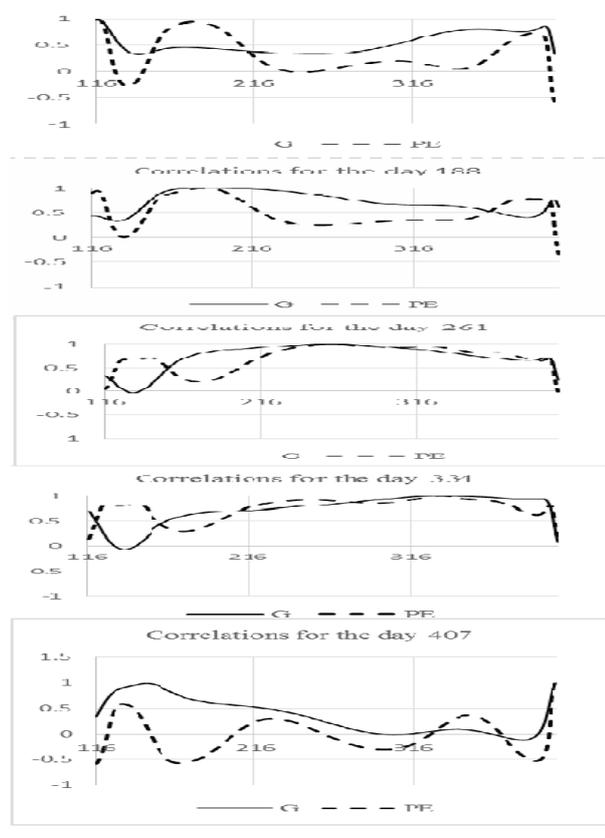


Figure 6 Genetic (solid line) and permanent environmental (dash line) correlations for 5 values of the control variable, corresponding to lowest and highest values (d 116 and d 407, respectively) and 3 approximately equidistant intermediate values (d 188, d 261 and d 334)

The range of permanent environmental correlations for five control variables including 116, 188, 261, 334 and 407 were -0.492 to 0.99, -0.371 to 0.99, -0.15 to 0.99, 0.146 to 0.99 and -0.594 to 0.99, respectively. The permanent environmental correlations were dominantly positive for the first three points and were all positive for d 334. Permanent environmental correlations for three middle control variables including 188, 261 and 334 had wild oscillations in peak of the first egg production cycle. As reported by Wolc *et al.* (2009) the range of permanent environmental correlations for partial egg production were from -0.12 to 0.99 which in the lower bound was different from the estimations in the current study. Wild oscillations for correlations between the first (d 116) and the last (d 407) control variable lasted for the whole egg production period. As Arango *et al.* (2004) mentioned the apparent weakness of the model for fitting the trait might be due to 1) a lack of mathematical flexibility to model the trait adequately with orthogonal polynomials; 2) the small number of records at different control variables; and 3) the general property of the estimation of regression coefficients. Rawlings *et al.* (1998) indicated that even with polynomial models of higher order and associated smaller residual sum of squares, in specific ages might exhibit wild oscillations.

CONCLUSION

Egg production curve of Iranian native fowl is not a typical egg production curve. Using random regression model appears promising and could be recommended for estimation of variance components of CS. Although none of the studies on random regression modeling used CS as repeated measurements, the results obtained by the current study proposed that the model composed of the second order polynomial for ASM and 8th order polynomial for additive genetic and permanent environmental effect was the most suitable for estimation of variance components in the current dataset. Selection applied in the CS can result in gains for adjacent ages. However, as the distance in time between consecutive ages increase, the selection becomes less efficient. Because the genetic correlations for d261 was dominantly greater than 0.5, the selection made on the basis of d242, where highest heritability was estimated, is useful.

ACKNOWLEDGEMENT

Our sincere gratitude goes for the staff of the Native Fowl Breeding center of the provincial Organization of Agriculture for collecting the data.

REFERENCES

Akbas Y., Takma Ç. and Yaylak E. (2004). Genetic parameters for

- quail body weights using a random regression model. *Southern African J. Anim. Sci.* **34(2)**, 104-109.
- Akbas Y., Ünver Y., Oguz I. and Altan O. (2002). Estimation of genetic parameters for clutch traits in laying hens. Proc. 7th World Cong. Genet. Appl. Livest. Prod., Montpellier, France.
- Anang A., Mielenz N. and Schüler L. (2002). Monthly model for genetic evaluation of laying hens. II. Random regression. *Br. Poult. Sci.* **43(3)**, 384-390.
- Anang A., Mielenz N., Schüler L. and Preisinger D.R. (2001). The use of monthly egg production records for genetic evaluation of laying hens. *J. Ilmu. Ternak dan Vet.* **6(4)**, 230-234.
- Arango J., Cundiff L.V. and Van Vleck L.D. (2004). Covariance functions and random regression models for cow weight in beef cattle. *J. Anim. Sci.* **82(1)**, 54-67.
- Crow F. and Dove W.F. (2000). Perspectives on Genetics: Anecdotal, Historical and Critical Commentaries. The University of Wisconsin Press. 2537 Daniels Street, Madison, Wisconsin.
- Farzin N., Vaez Torshizi R., Kashan N. and Gerami A. (2010). Estimation of genetic and phenotypic correlations between monthly and cumulative egg productions in a commercial broiler female line. *Glob. Vet.* **5(3)**, 164-167.
- Grossman M., Grossman T.N. and Koops W.J. (2000). A model for persistency of egg production. *Poult. Sci.* **79**, 1715-1724.
- Jensen J. (2001). Genetic evaluation of dairy cattle using test-day models. *J. Dairy Sci.* **84**, 2803-2812.
- Kirkpatrick M., Lofsvold D. and Bulmer M. (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics.* **124**, 979-993.
- Komprej A., Malovrh Š., Gorjanc G., Kompan D. and Kovac M. (2013). Genetic and environmental parameters estimation for milk traits in Slovenian dairy sheep using random regression mode. *Czech J. Anim. Sci.* **58(3)**, 125-135.
- Koops W.J. and Grossman M. (1992). Characterization of poultry egg production using a multiphasic approach. *Poult. Sci.* **71**, 399-405.
- Kranis A., Su G., Sorensen D. and Woolliams J.A. (2007). The application of random regression models in the genetic analysis of monthly egg production in turkeys and a comparison with alternative longitudinal models. *Poult. Sci.* **86(3)**, 470-475.
- Lukovic Z., Uremovic M., Konjacic M., Uremovic Z. and Vincek D. (2007). Genetic parameters for litter size in pigs using random regression model. *Asian-australs J. Anim. Sci.* **20(2)**, 160-165.
- Lukovic Z., Malovrh S., Gorjanc G. and Kovac M. (2004). A random regression model in analysis of litter size in pigs. *Southern African J. Anim. Sci.* **34(4)**, 241-248.
- Lukovic Z., Malovrh S., Gorjanc G. and Kovac M. (2003). Genetic parameters for number of piglets born alive using random regression model. *Agric. Cons. Sci.* **68(2)**, 105-108.
- Meyer K. (2006). WOMBAT. University of New England. Australia.
- Meyer K. (1998). Estimating covariance functions for longitudinal data using a random regression model. *Genet. Sel. Evol.* **30**, 221-240.
- Minvielle F., Kayang B.B., Inoue-Murayama M., Miwa M., Vignal A., Gourichon D., Neau A., Monvoisin L. and Ito S.I. (2006). Search for QTL affecting the shape of the egg laying curve of the Japanese quail. *BMC Genet.* **7**, 26-31.
- Miyoshi S., Luc K.M. and Kuchida K. (1996). Application of non-linear models to egg production curves in chickens. *Japans Poult. Sci.* **33**, 178-184.
- Pool M.H. and Meuwissen T.H.E. (1999). Prediction of daily milk yields from a limited number of test days using test day models. *J. Dairy Sci.* **82**, 1555-1564.
- Rawling J.O., Pantula S.G. and Dickey D.A. (1998). Applied Regression Analysis: A Research Tool. Springer Texts in Statistics. Springer-Verlag, Inc., New York.
- Rowland H.D. and Weller S.G. (1996). The Gist of Genetics: Guide to Learning and Review. Jones and Barlett Publishers, London, UK.
- Strabel T., Szyda J., Ptak E. and Jamrozik J. (2005). Comparison random regression test-day model for polish black and white cattle. *J. Dairy Sci.* **88**, 3688-3699.
- Van der Werf J. (2001). Random Regression in Animal Breeding. Course Notes. University of New England. Armindale. Australia.
- Venturini G.C., Grossi D.A., Ramos S.B., Cruz V.A.R., Souza C.G., Ledur M.C., Faro L.El., Schmidt G.S. and Munari D.P. (2012). Estimation of genetic parameters for partial egg production periods by means of random regression models. *Genet. Mol. Res.* **11(3)**, 1819-1829.
- Wolc A., Arango J., Settar P., Fulton J.E., O'Sullivan N.P., Preisinger R., Fernando R., Garrick D.J. and Dekkers J.C.M. (2013). Analysis of egg production in layer chickens using random regression model with genomic relationships. *Poult. Sci.* **92(6)**, 1486-1491.
- Wolc A., Arango J., Settar P., O'Sullivan N.P. and Dekkers J.C.M. (2011). Evaluation of egg production in layers using random regression models. *Poult. Sci.* **90(1)**, 30-34.
- Wolc A., Bednarczyk M., Lisowski M. and Szwaczkowski T. (2010). Genetic relationships among time of egg formation, clutch traits and traditional selection traits in laying hens. *J. Anim. Feed Sci.* **19**, 452-459.
- Wolc A. and Szwaczkowski T. (2009). Estimation of genetic parameters for monthly egg production in laying hens based on random regression models. *J. Appl. Genet.* **50(1)**, 41-46.
- Wolc A., Lisowski M. and Szwaczkowski T. (2007). Heritability of egg production in laying hens under cumulative, multitrait and repeated measurement animal models. *Czech J. Anim. Sci.* **52(8)**, 254-259.