



A Comparative Study of Simple Regression Models to Estimate Fibre Length Growth in Chios Sheep from Common Meteorological Variables

ARISTIDIS MATSOUKIS*, AIKATERINI CHRONOPOULOU-SERELI
and GEORGE STRATAKOS

Department of Crop Science, Agricultural School of Agricultural Production, Infrastructure and Environment, Agricultural University of Athens, Athens, Greece.

Abstract

Chios sheep is a promising sheep breed, with wool, one of its products, to be of special interest to genetic improvement programs. Recently, it has been reported significant linear correlation between the fibre length growth (FLG) of Chios sheep, an important component of its wool quality, and each of the meteorological variables air temperature (T) and sunshine (SUNS), but nothing is known about the prediction of FLG from T and SUNS. Thus, this work aims to investigate the effectiveness of five simple regression models (linear, quadratic, cubic, logarithmic and inverse), concerning the aforementioned prediction, using visual examination and two widely accepted statistical measures, the adjusted coefficient of determination (R^2_{adj}) and the root mean square error (RMSE). Results showed that the applied nonlinear regression models were characterized by higher R^2_{adj} and lower RMSE in comparison to the linear one, irrespective of input variable. The inverse model presented the greatest effectiveness to predict FLG from T and SUNS, separately (maximum R^2_{adj} and minimum RMSE), followed by the logarithmic and the linear ones, under visual examination and applied statistical measures. Air temperature was superior to SUNS in all cases (higher R^2_{adj} and lower RMSE), when comparing the regression models of the same type to check their effectiveness for predicting FLG. The findings of our study could be a decisive step towards a better exploitation of the examined meteorological variables for the sustainable production of Chios sheep.



Article History

Received: 06 October 2020

Accepted: 14 December 2020

Keywords:

Air Temperature;
Chios Sheep;
Coefficient Of Determination;
Fibre Length Growth;
Linear Models;
Nonlinear Regression;
Root Mean Square Error;
Sunshine.

CONTACT Aristidis Matsoukis ✉ armatsoukis@aua.gr 📍 Department of Crop Science, Agricultural School of Agricultural Production, Infrastructure and Environment, Agricultural University of Athens, Iera Odos 75, 11855, Athens, Greece.



© 2020 The Author(s). Published by Enviro Research Publishers.

This is an Open Access article licensed under a Creative Commons license: Attribution 4.0 International (CC-BY).

Doi: <http://dx.doi.org/10.12944/CARJ.8.3.04>

Introduction

Chios sheep is one of the most commonly reared indigenous sheep breeds in Greece¹ with very good performance in milk production.^{2,3} This breed has been participated in many genetic improvement programs^{2,4,5}, showing great potential. One of the products of Chios sheep in which genetic improvement programs have been focused recently is wool⁵ which is irreplaceable for the production of high quality fabrics.^{2,6} An important component of the wool quality is the length of fibre⁷ but, up to now, little research has been conducted on this component of Chios sheep. Matsoukis *et al.*,² recently reported a seasonal variation of wool fibre length of this breed in relation to important meteorological variables.

Air temperature (T) is a meteorological variable of primary importance for farm animals.^{8,9} It impacts considerably their physiological functions⁸ and probably the fibre length growth (FLG) of Chios sheep taking into account the reported significant correlation between T and FLG in this breed. A similar significant correlation has been reported between the aforementioned variable and sunshine (SUNS)², another common meteorological variable, which has also been reported to affect the postpartum ovarian activity of the Murrah breed of buffalo (*Bubalus bubalis* L.) in India.¹⁰

Regression analysis is a statistical tool which helps evaluating which variables are capable of predicting a single, continuous dependent variable.^{11,12} Regression technique may be divided into two broad categories, that is, linear regression models and nonlinear regression models. Linear regression models are the simplest ones and commonly used in both data analysis and as a part of the learning process.¹³ On the other hand, the more realistic models are often the nonlinear ones. A characteristic example of these models is many growth models.¹⁴

Literature, to our knowledge, presents a gap regarding the prediction of FLG of Chios sheep from T and SUNS. Thus, a need arises to predict Chios sheep FLG from these meteorological variables, using various regression models, linear and nonlinear ones, and compare them for their effectiveness. The basic hypothesis for our current work is that the examined nonlinear models, that is, quadratic, cubic, logarithmic and inverse ones,

better predict FLG from T or SUNS than the linear one, taking into consideration that nonlinear models seem to approach better reality.¹⁴

Materials and Methods

Research was carried out in the farm of Artificial Insemination Center of Karditsa Municipality (39°21'18''N, 21°54'19''E), Periphery of Thessaly, Greece, for a two-year period. Six rams (*Ovis aries* L.) of Chios sheep breed, all about six years old, living in the aforementioned farm, were the experimental material. The examined variables were FLG and two meteorological ones, T and SUNS, concerning the experimental environment of the animals. Details regarding the measuring instruments, their good operation, time of measurements and recording, ways of measurement, estimation and processing of FLG, T and SUNS were reported in Matsoukis *et al.*² Briefly, fibre length measurements were conducted manually once a month with the aid of a ruler. From these measurements, FLG was estimated as their difference between adjacent months. Air temperature was recorded continually every 6 min by appropriate sensors (Hobo type Pro, H08-032-08, USA, accuracy $\pm 0.2^\circ\text{C}$ at 25°C) connected to data loggers and sunshine data were kindly provided by the meteorological station of the Tobacco Station of Agricultural Research, adjacent to the farm.² For our study, mean values of all variables were estimated for each month, for the whole studied period, and were used in the statistical analysis.

Specifically, in order to detect the possible response function of FLG in Chios sheep to each one of the examined meteorological variables, the first step of our statistical analysis was the test of easily available simple regression models. Thus, we applied five models, one linear and four nonlinear, that is, the quadratic, cubic, logarithmic and inverse ones^{15,16}, defined by the following equations:

$$y = a + b_1 \cdot x \text{ (linear)} \quad \dots(1)$$

$$y = a + b_1 \cdot x + b_2 \cdot x^2 \text{ (quadratic)} \quad \dots(2)$$

$$y = a + b_1 \cdot x + b_2 \cdot x^2 + b_3 \cdot x^3 \text{ (cubic)} \quad \dots(3)$$

$$y = a + (b_1 \cdot \ln x) \text{ (logarithmic)} \quad \dots(4)$$

$$y = a + (b_1/x) \text{ (inverse)} \quad \dots(5)$$

where y : dependent variable (FLG), x : independent variable (separately T and SUNS), a : Y-axis intercept, b_1, b_2, b_3 : regression coefficients and \ln : natural logarithm.

Then, to evaluate the performance of the examined models, specifically, their goodness of fit (quality of prediction), leading to their comparison, a careful visual examination of the fitted regression lines and two well acceptable statistical measures were used; the adjusted coefficient of determination (R^2_{adj})^{17,18} and the root mean square error (RMSE).^{17,19} The R^2_{adj} , an adjustment of the coefficient of determination, explains how well a model fits a data. It indicates the proportion of total variation in the response (dependent) variable that is explained by the predictor (independent) variables. The bigger the R^2_{adj} the better the model.¹⁹ The RMSE, a sort of generalized standard deviation²⁰, is used to measure the differences between the values predicted by a model and the observed values.¹⁹ It is considered as one of the most important criteria to compare the suitability of regression models. The smaller its value, the better the model fits the data.²⁰ The statistics was performed using MS Excel 2007^{21,22} and IBM SPSS Statistics 23 with results to be considered significant at $P \leq 0.05$.

Results and Discussion

The monthly averages concerning FLG of Chios sheep ranged from 1.1 cm (October) to 2.5 cm (January) with 1.68 ± 0.07 to be their mean \pm standard error of the mean (SE). As regards T and SUNS, their monthly averages ranged from 3.3°C (January) to 28.6°C (July) and 3700 h (December) to 19585 h (July), respectively, with their means \pm SE to take the values $16.2 \pm 1.7^\circ\text{C}$ (T) and 11016.1 ± 993.2 h (SUNS).

The statistical results of the tested regression models with FLG regressed on T and SUNS, separately, are shown in Tables 1 and 2, respectively. For the examined linear and nonlinear models, significant R^2_{adj} values were confirmed. These values ranged from 0.298 to 0.544 with regard to T and 0.200 to 0.292 with regard to SUNS. The higher and the lower R^2_{adj} values were estimated for the inverse and linear models, respectively, for both T and SUNS. The RMSE ranged from 0.226 to 0.281 and 0.282

to 0.300, with regard to the model inputs T and SUNS, respectively. The lower and the higher values of RMSE were confirmed in the cases of inverse regression model and linear one, respectively, irrespective of the aforementioned variables.

The goodness of fit of the tested regression models, from a purely statistical point of view, taking into account that the best and the worst models were characterized by the combinations of maximum R^2_{adj} with minimum RMSE and minimum R^2_{adj} with maximum RMSE, respectively, followed the descending orders Inverse>Cubic>Quadratic>Lo-garithmic>Line-ar for T as input variable (Table 1) and Inverse>Quadratic>Lo-garithmic>Cubic>Line-ar for SUNS input (Table 2). When comparing the regression models of the same type to investigate the effectiveness of the independent variables (T and SUNS) for predicting FLG, T was superior to SUNS in all cases, due to its coincidence with higher R^2_{adj} and lower RMSE. The aforementioned superiority may be attributed, in general, to the noticeable role of T to the performance of domesticated farm animals²³, like sheep.²⁴

Although the cubic and quadratic models appeared a high rank, in general, regarding their effectiveness to predict FLG, taking into account the applied statistical criteria (Tables 1 and 2), the visual examination of the respective fitted regression lines (data not shown) led us to the ascertainment that these models are unsuitable to predict sufficiently the FLG of Chios sheep. The quadratic and cubic models were characterized by one hump (an actual shape of asymmetric U) and two humps (one facing upward and the other down), respectively.²⁵ These humps had no physical meaning, taking also into consideration the growth pattern of sheep fibre length in relation to the examined meteorological variables.^{2,26,27,28} Thus, finally, under the given data and time period, the more suitable and consequently recommended regression model to predict FLG was the inverse model, followed by the logarithmic and the linear one (last in the rank order list), independently of the studied predictors (T and SUNS). Nevertheless, a greater time period is required for a more precise estimation of FLG, based on T and SUNS, for Chios sheep.

Table 1: Regression parameters of the tested regression models concerning air temperature (independent variable) and fibre length growth (dependent variable) in Chios sheep

Regression model	a	b ₁	b ₂	b ₃	R ² _{adj}	RMSE
Linear	2.053	-0.023	-	-	0.298**	0.281
Quadratic	2.608	-0.113	0.003	-	0.460**	0.246
Cubic	3.378	-0.318	0.017	0.0001	0.537***	0.228
Logarithmic	2.635	-0.364	-	-	0.438***	0.251
Inverse	1.337	3.849	-	-	0.544***	0.226

a: Y-axis intercept, b₁, b₂, b₃: Regression coefficients, ***, **: Significance at P≤0.001 and P≤0.01, respectively, RMSE: Root mean square error.

Table 2: Regression parameters of the tested regression models concerning sunshine (independent variable) and fibre length growth (dependent variable) in Chios sheep

Regression model	a	b ₁	b ₂	b ₃	R ² _{adj}	RMSE
Linear	2.046	-3.334·10 ⁻⁵	-	-	0.200*	0.300
Quadratic	2.681	0.0002	5.183·10 ⁻⁹	-	0.287*	0.283
Cubic	3.100	0.0001	1.879·10 ⁻⁸	-3.914·10 ⁻¹³	0.262*	0.288
Logarithmic	5.285	-0.392	-	-	0.269**	0.286
Inverse	1.301	3386.846	-	-	0.292**	0.282

a: Y-axis intercept, b₁, b₂, b₃: Regression coefficients, **, *: Significance at P≤0.01 and P≤0.05, respectively, RMSE: Root mean square error.

Conclusion

In conclusion, the examined simple nonlinear regression models (inverse, logarithmic, cubic, quadratic) were characterized by higher R²_{adj} and lower RMSE in comparison to the linear model. The inverse model showed the greatest effectiveness in predicting FLG from T and SUNS, separately (maximum R²_{adj} and minimum RMSE), followed by the logarithmic and the linear ones, under visual examination and applied statistical criteria. The comparison of the regression models of the same type to check the effectiveness of T and SUNS for predicting FLG of Chios sheep, showed that T was superior to SUNS in all cases (higher R²_{adj} and lower RMSE). The findings of our study could be helpful in establishing sophisticated plans towards a better

exploitation of the meteorological variables for the sustainable production of this interesting sheep breed.

Acknowledgements

Thanks are due to State Scholarships Foundation of Greece which partly funded the experimental part of this study.

Funding

The author(s) received no additional financial support for the research, authorship, and/or publication of this article.

Conflict of Interest

The authors declare that there is no conflict of

interest.

References

1. Argyriadou A., Gelasakis A., Banos G., Arsenos G. Genetic improvement of indigenous Greek sheep and goat breeds. *J Hellenic Vet Med Soc.* 2020; 71(1): 2063-2072.
2. Matsoukis A., Stratakos G., Chronopoulou-Sereli A., Tsiros I. Seasonal variation of wool fibre length in Karagouniko and Chios sheep in relation to meteorological factors. *Emir J Food Agric.* 2019; 31(10): 788-793.
3. Rogdakis E. Indigenous Breeds of Sheep. Description- Phylogeny-Genetic Improvement-Preservation. Athens, Greece: Agrotipos Publications; 2002.
4. Gelasakis A.I., Arsenos G., Hickford J., Zhou H., Psifidi A., Valergakis G.E., Banos G. Polymorphism of the MHC-DQA2 gene in the Chios dairy sheep population and its association with footrot. *Livest Sci.* 2013; 153: 56-59.
5. Yardibi, H., Gürsel, F.E., Ates, A., Akis, I., Hacıhasanoğlu, N., Özdem Öztapak, K. Polymorphism of the Kap 1.1, Kap 1.3 and K33 genes in Chios, Kivircik and Awassi. *Kafkas Univ Vet Fak Derg.* 2015; 21 (4): 535-538.
6. Zygoiannis D. Sheep Farming. 3rd ed. Thessaloniki, Greece: Modern Education Publications; 2014.
7. Mendel C. Practical Sheep Farming. Rearing-Reproduction- Breeds-Hygiene-Products-Legislation. Athens, Greece: Vasdekis Publications; 2010.
8. Bengtsson L.P., Whitaker J.H. Farm structures in tropical climates. Nairobi, Kenya: Food and Agriculture Organization of the United Nations (FAO) and the Information Network on Post-Harvest Operations (INPhO); 1988.
9. Ames D.R., Curtis S.E., Hahn G.L., McDowell R.E., Polin D., Young B.A. Effect of environment on nutrient requirements of domestic animals. Washington, D.C., USA: National Academies Press; 1981.
10. Abayawansa W.D., Prabhakar S., Singh A.K., Brar P.S. Effect of climatic changes on reproductive performance of Murrah buffaloes in Punjab: A retrospective analysis. *Indian J Anim Sci.* 2011; 81(4): 334-339.
11. Kleinbaum D.G., Kupper L.L., Muller K.E., Nizam A. Applied regression analysis and other multivariate methods. 3rd edition. USA: Duxbury Press; 1998.
12. Moore D.S., Notz W.I., Fligner M.A. The basic practice of Statistics. 6th edition. New York, USA: W.H. Freeman and Company; 2013.
13. Rencher A.C., Schaalje GB. Linear Models in Statistics. 2nd edition. Hoboken, New Jersey, USA: John Wiley & Sons Inc.; 2008.
14. Rawlings J.O., Pantula S.G., Dickey D.A. Applied regression analysis, a research tool. Springer texts in statistics. 2nd edition. New York, USA: Springer-Verlag; 1998.
15. Ates, Akis, Hacıhasanoğlu with Ateş, Akiş, Hacıhasanoğlu, respectively.
16. Curve estimation models. IBM Knowledge Center. https://www.ibm.com/support/knowledge center/SSLVMB_23.0.0/spss/base/curve_estimation_models.html. Accessed 21 September 2020.
17. Ghavi Hossein-Zadeh N., Golshani M. Comparison of non-linear models to describe growth of Iranian Guilan sheep. *Rev Colomb Cienc Pecu.* 2016; 29:199-209.
18. Kaplan S., Gürcan E.K. Comparison of growth curves using non-linear regression function in Japanese quail. *J Appl Anim Res.* 2018; 46(1):112-117.
19. Zhai H. Linear and non-linear regression: application to competitor's gasoline volume estimation. Lup students papers, Lund University Libraries, Lund University. <https://lup.lub.lu.se/student-papers/search/publication/5034934>. Published 29 January 2015. Accessed 25 September 2020.
20. Ghavi Hossein-Zadeh N. Comparison of non-linear models to describe the lactation curves of milk yield and composition in Iranian Holsteins. *J Agric Sci.* 2014; 152:309-324.
21. Matsoukis A., Chronopoulou-Sereli A., Chronopoulos I. Bioclimatic conditions in relation to shading in a glasshouse: The case study of *Lantana camara* L. cultivation in summer. *Curr Agri Res J.* 2016; 4(1): 47-53.

22. Matsoukis A., Kamoutsis A., Chronopoulou-Sereli A. A note on the flowering of *Ajuga orientalis* L. in relation to air temperature in Mount Aenos (Cephalonia, Greece). *Curr Agri Res J.* 2018; 6(3): 261-267.
23. Chronopoulou-Sereli A., Flocas A. *Lessons of Agricultural Meteorology and Climatology*. Thessaloniki, Greece: Ziti Publications; 2010.
24. Sawyer G., Narayan E.J. A review on the influence of climate change on sheep reproduction. *Comparative Endocrinology of Animals*. IntechOpen; 2019: 1-21. <https://www.intechopen.com/books/comparative-endocrinology-of-animals/a-review-on-the-influence-of-climate-change-on-sheep-reproduction>. Accessed 4 December 2020.
25. Grace-Martin K. *Regression Models: How do you know you need a polynomial? The analysis factor*. <https://www.theanalysisfactor.com/regression-modelshow-do-you-know-you-need-a-polynomial/>. Accessed 29 September 2020.
26. Schlink A.C., Mata G., Lea J., Ritchie A.J.M. Seasonal variation in fibre diameter and length growth rate in high and low staple strength Merino wethers. *Proc Aust Soc Anim Prod.* 1996; 21: 356.
27. Elsherbiny A.A., Eloksh H.A., Elsheikh A.S., Khalil M.H. Effect of light and temperature on wool growth. *J Agric Sci.* 1978; 90: 329-334.
28. Schlink A.C., Mata G., Lea J.M., Ritchie A.J.M. Seasonal variation in fibre diameter and length in wool of grazing Merino sheep with low or high staple strength. *Aust J Exp Agric.* 1999; 39: 507-517.