APPLICATION OF WILMINK CURVE AND ARTIFICIAL NEURAL NETWORK TO PREDICT MILK YIELD IN COWS

Wilhelm Grzesiak, Piotr Błaszczyk, Piotr Sablik

Department of Ruminant Science, Agricultural University of Szczecin, Poland

In 1987, Wilmink (1987) proposed a function for analysing lactation yield in dairy cows that has since been widely used by many authors for either bovine lactation curve modelling (Ptak and Żarnecki 1988; Schaeffer et al. 2000; Ptak and Frącz 2002) or cattle genetic evaluation (the Test-Day Model).

Artificial neural networks (ANN) are known for their ability to learn, which can substitute programming. They have been successfully applied in a number of livestock management-related problems, e.g. housing environment control on a swine farm (Korthals et al. 1994), hoof cracking modelling in horses (Suchorski-Tremblay et al. 2001), dairy cow standard lactation yield prediction (Lacroix et al. 1995), mastitis monitoring (Yang et al. 1999), or detecting oestrus or mammary gland inflammations in cows (DeMol and Woldt 2001). It should be noted, however, that many attempts to apply ANN have not yielded the expected results, which has not always been published (Tadeusiewicz 1998).

The aim of this study was an attempt to determine the ability of a neural network model to predict full lactation milk yield in cows managed in a given barn, basing on daily yields (RZ), and a comparison of the quality of the neural network model (NN) with a regression model (WL), as well as the predictions by both models compared with official milk recording system SYMLEK (SL) predictions.

Material and methods

Basing on daily milk yields obtained from first lactation (53.388), Wilmink (WL) model parameters were estimated and an artificial neural network was trained and verified. Remaining daily yields (14.234) were used to test the prognostic quality of the studied models. The cows were divided into two genetic groups (less or equal to 75% HF and 75.1-100% HF), two calving season groups (autumn-winter and spring-summer calving), as well as two groups according to age at calving (18-30 months and 31-46 months for first lactation, 28-39 months and 40-64 months for second lactation, and 37-53 months and 54-89 month for third lactation), which resulted in a total number of eight curve equations for first lactation, $y = a + bt + ce^{-0.05t}$ (a,b,c – parameters; t – days of milk)

The NN model was based on the following predictor variables: x_1 – HF genotype percentage, x_2 – age at calving in months, x_3 – month of calving (October encoded as 1, November as 2, December as 3, January as 4, and so forth, until September encoded as 12), x_4 – day in milk after calving, x_5 – lactation number. Selection of the predictors was based on available data, so that the NN processed the same information as the estimated regression models. The dependent variable in the NN model was the actual milk yield on a given day.

From among the available network models a perceptron was selected with two hidden layers which had 10 and 6 neurons in, respectively, the first and the second layer. The parameters describing the quality of the network, SD_{ratio} (the ratio of the prediction error standard deviation to the original output data standard deviation) and r (Pearson's linear correlation coefficient between the NN input data and output results), were the best. Input and output data sets were converted using minimax linear scaling.

The network training was based on the error back-propagation algorithm (Osowski 2000). This process was performed for 30,000 epochs (learning steps), which involved a single presentation of all the cases of the training data set and, based on it, modification of the network parameters. A decreasing learning rate (η=0.9 down to η= 0.5) and the momentum coefficient (α=0.5) were adopted for each tested network. Sigmoid function was applied as the activation function.

The quality of the network and regression models was determined with: coefficient of determination (R^2); model global relative approximation error (RAE), and squared root mean square error (RMS). In order to evaluate the prediction properties of the analysed models, we used the following: Pearson's linear correlation coefficient (r) between the model-predicted values and the actual yields, average relative prediction error (ψ), and Theil's inequality coefficient, I^2 (Cieślak 1997).

Results and discussion

The neural network model demonstrated better quality parameters compared to Wilmink models: $REA_{NN} = 0.14$, $RAE_{WL} = 0.17$, $RMS_{NN} = 2.65$, $RMS_{WL} = 3.23$, $R^2_{NN} = 0.75$, and $R^2_{WL} = 0.48$. Also the prediction indices were better for the NN model ($\psi_{NN} = 13.01$, $\psi_{WL} = 16.16$; $r_{NN} = 0.85$, $r_{WL} = 0.62$; $I^2_{NN} = 0.023$, $I^2_{NN} = 0.031$).

Selecting the best NN model in terms of lactation yield prediction, we paid special attention to the magnitude of the SD_{ratio} coefficient and the coefficient of correlation between the input values and the predicted yields. In order to successfully train the network, an appropriate quality of data representing the given problem is needed, which has been raised by Lacroix et al. (1997).

Wilmink models had on average R2 = 0.48. Wilmink (1987) achieved 0.97, Catillo et al. (2002) reported 0.99, while Freeze and Richards (1992) obtained a more similar value to ours, i.e. 0.51. The qualitative parameters obtained for the neural network were better compared to the Wilmink curves.

Predictions by the neural network were more accurate in the initial stage of lactation (Table 1). Rowlands et al. (1982) obtained a similar prediction patterns.

Table 1. Actual cumulative yields by lactation stage and predictions estimated by SYMLEK and analysed models and correlations between the predictions and actual data

Days	RZ	SL	*WL*	*NN*	RZ– SL	RZ– WL	RZ– NN	r_{SL}	r_{WL}	r_{NN}
5-20	303	235	314	299	68	-11	4	*0,8962*	*0,9277*	*0,9475*
21-40	731	639	732	726	92	-1	5	*0,9095*	*0,9481*	*0,9566*
41-60	1157	1069	1161	1147	88	-4	10	*0,9005*	*0,9622*	*0,9601*
61-80	1564	1466	1560	1551	98	4	13	*0,8913*	*0,9207*	*0,9461*
81-100	1950	1993	1983	1946	-43	-33	4	*0,9305*	*0,9631*	*0,9174*
101-120	2320	2307	2355	2331	13	-35	-11	*0,8664*	*0,8776*	*0,9126*
121-140	2686	2666	2737	2708	20	-51	-22	*0,8756*	*0,8793*	*0,8989*
141-160	3051	3108	3107	3073	-57	-56	-22	*0,8165*	*0,9457*	*0,9701*
161-180	3413	3492	3450	3427	-79	-37	-14	*0,7631*	*0,9449*	*0,9408*
181-200	3766	3917	3797	3770	-151	-31	-4	*0,6961*	*0,8378*	*0,8973*
201-220	4064	4126	4129	4100	-62	-65	-36	*0,8723*	*0,8836*	*0,8988*
221-240	4367	4589	4462	4418	-222	-95	-51	*0,8437*	*0,9152*	*0,8897*
241-260	4671	4890	4783	4727	-219	-112	-56	*0,8597*	*0,8639*	*0,8764*
261-280	4947	5167	5065	5023	-220	-118	-76	*0,8079*	*0,8445*	*0,8858*
281-305	5286	5600	5402	5379	-314	-116	-93	*0,8975*	*0,9343*	*0,9575*
Razem					-988	-761	-349	*0,8565*	*0,9112*	*0,9315*
	Absolutely differences				*1745*	*767*	*420*			

Lacroix et al. (1995) observed that prediction abilities differ depending on the type of network and/or lactation stage. No doubt, the cows with a more flat lactation curve are more desired, since they demonstrate stronger endurance and higher lactation yields (Tekereli et al. 2000).

It should be stressed that lactation yield predictions by the SYMLEK system were much worse than the predictions obtained from the neural network model or from linear models as well. A conversion into 305-d lactation was not that clear-cut, however, since the particular final predictions did not differ much from the actual yields (Table 1 – in bold). Coefficients of correlation did not differ significantly, hence we have presumed that the fitness of predictions with the actual data for the cows was similar (r = 0.90). Wilmink (1987) stated that the correlation between predictions and test yields were 0.86. Similar correlations (0.84-0.97) were found by Kominakis et al. (2000) for sheep milk yields.

References

1. Catillo G., Macciotta N.P., CarrettaA., Cappio-Borlino A. 2002. J. Dairy Sci. 85(5): 1298-1306.
2. Cieślak M. 1997. Prognozowanie gospodarcze. Metody i zastosowania. PWN, Warszawa.
3. De Mol R.M., Woldt W.E. 2001. J. Dairy Sci. 83(5): 1094-1103.
4. Freeze B.S., Richards T.J. 1992. J. Dairy Sci. 75:2984-2989.
5. Kominakis A.P, Abas Z., Maltaris I., Rogdakis E. 2002. Comp.Electron. Agric.35:35-48.

6. Korthals R.L., Hahn G.L., Nienaber J.A. 1994. Trans. ASAE 37(4):1295-1299.
7. Lacroix R., Salehi F., Yang X.Z., Wade K.M. 1997. Trans. ASAE 40(3):839-846.
8. Lacroix R., Wade K.M., Kok R., Hayes J.F. 1995. ASAE 38(5):1573-1579.
9. Osowski S. 1997. Sieci neuronowe ujęciu algorytmicznym. WNT, Warszawa.
10. Ptak E., Frącz A. 2002. Rocz. Nauk. Zootech. 15: 31- 36.
11. Ptak E., Żarnecki A. 1998. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia 1998. Vol. 13. [b.w.]: 335-338.
12. Rowlands G.J., Lucey S., Russell A.M. 1982. Anim. Prod. 35: 135-144.
13. Schaeffer L.R., Jamrozik J., Kistemaker G.J., Doormaal B.J. 2000.. J. Dairy Sci. 83:1135-1144.
14. Suchorski-TremblayA.M., Kok R., Thompson J.J.2001.Canad.Biosyst.Eng.43:7.15-7.22.
15. Tadeusiewicz R. 1998. Mat. Sem."NeuroMet' 98", Kraków 1998,7-25.
16. Tekerli M., Akinci Z., Dogan I., Akcan A. 2000. J. Dairy Sci. 83: 1381-1386.
17. Wilmink J. B. M. 1987. Livest. Prod. Sci. 16: 335-348.
18. Yang X.Z., Lacroix R., Wade K.M. 1999. Trans. ASAE 42(4): 1063-1071.

Summary

The neural network model has met our expectations in terms of prediction quality. Its application for lactation curve analysis and for predicting 305-d lactation yields seem justified, since the quality parameters of neural networks were better that those for the Wilmink model. It was found that daily-yield-based predictions by regression models and neural network were more accurate that those by the official milk recording system (SYMLEK). A properly prepared (trained) neural network may generate yield predictions for individual cows or group of cows, while application of mathematical models requires that a new equation be developed each time the group of cow changes.

SUITABILITY OF DOUBLE HYBRIDS OF BILGORAJSKIE AND WHITE ITALIAN GEESE FOR OAT FATTENING

Helena Puchajda, Krzysztof Pudyszak

Department of Poultry Science, University of Warmia and Mazury in Olsztyn

Introduction

In Poland, the production of geese is based exclusively on White Italian geese that were imported from Denmark in 1962. These geese acclimatized quickly to new environmental conditions and became competitive with domestic geese because of high and versatile usability. This caused a gradual displacement of domestic races and lines of geese, among which ten groups make now a genetical resource [2, 7]. The groups of mainatined geese are used in research to produce commercial hybrids, among others [3, 6]. One of the domestic goose lines used in many studies is Bilgorajskie geese, being maintained as a closed flock The observations made about this flock for many years indicate their high usability. These geese are characterized by good health conditions, good feed utilization, high slaughter value, and also a considerable con-