Available online at www.sciencedirect.com

ScienceDirect

Computers and Electronics in Agriculture 54 (2006) 69–83

Computers and electronics in agriculture

www.elsevier.com/locate/compag

Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs)

Wilhelm Grzesiak [a,*], Piotr Błaszczyk [a], René Lacroix [b]

[a] *Laboratory of Biostatistics, Agricultural University of Szczecin, Szczecin, Poland*
[b] *McGill University, Montreal, Quebec, Canada*

Received 20 June 2005; received in revised form 29 August 2006; accepted 31 August 2006

Abstract

The study is focused on the capability of artificial neural networks to forecast milk yield for both full and standardised lactations. We used a dataset of 108,931 daily milk yields (dataset A) collected from three lactations of dairy cows managed in a production farm. Using the actual data on daily milk yields and the data recorded on official milk recording test days, a number of neural networks were designed and parameters of Wood's model were estimated. The quality of each network and regression model was measured using coefficients of determination, relative approximation errors (RAE), and root mean square errors (RMS). In order to test the prognostic parameters of the models, we randomly selected a subset of cows from the studied population, which produced in a dataset of 28,576 daily yields (dataset B). For those cows, daily and lactation yield forecasts were generated, which were next compared with their *actual* (observed) yield records and with the yields *calculated* by SYMLEK (ZETO Olsztyn Sp. z o.o., www.zeto.olsztyn.pl). The results have shown that the quality parameters of the designed neural networks were better than those of the regression model, for both the daily yields and test-day data (higher coefficients of determination and lower RAE and RMS). The prognostic parameters estimated for the forecasts of the neural networks were characterised by lower errors of prediction for both the daily yields and test-day data and exhibited higher coefficients of correlation between the predicted and the actual data (or the yields produced by SYMLEK). The predictions by the neural networks were more accurate than those by Wood's models. Furthermore, the predictions by both analysed models were closer to reality than the values estimated with the SYMLEK system. Application of neural networks does *not* require the data meeting the assumptions that must otherwise be met in a regression model. Large datasets are not needed to design a quite reliable neural network and, what is more, it is much easier to work with such a model than with a regression model.

Keywords: Dairy cows; Artificial neural networks; 305-d lactation; Wood's model; Daily yields

1. Introduction

The success of a dairy farm heavily depends upon decisions made at various stages of the production cycle. Since a tangle of genetic and environmental factors may often make correct decisions difficult, we need some tools to support breeding and management decisions.

* Corresponding author at: Laboratory of Biostatistics, Agricultural University of Szczecin, ul. Doktora Judyma 12, 71-460 Szczecin, Poland.
E-mail address: w.grzesiak@biot.ar.szczecin.pl (W. Grzesiak).

0168-1699/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2006.08.004

A mathematical formula describing the behaviour of milk yield during a lactation, referred to as *lactation curve*, may be a good instrument in dairy cows production monitoring. The shape of the curve provides the farmer with some valuable information. For example, cows with a very high peak yield are unable to take the amount of nutritional substances they need during the first stage of lactation, which is likely to lead to negative energy balance, reduced reproduction rates, and increased susceptibility to diseases (Jakobsen et al., 2002; Swalve, 2000). Conversely, cows with a flat lactation curve are more resistant to metabolic stress during the initial phase of lactation and have their energy demands more balanced, which in turn reduces feeding costs (Sölkner and Fuchs, 1987; Dekkers et al., 1998). Prediction of milk yield may also facilitate breeding decisions, e.g., selection, culling.

Wood (1967) proposed a *gamma* function for analysing performance of dairy cows as early as in 1967 and, although a range of other models emerged thereafter, this particular formula has been used up to now by many authors for modelling bovine lactation (Kellogg et al., 1977; Shanks et al., 1981; Schaeffer and Jamrozik, 1996; Leon-Velarde et al., 1995). The model has also been applied to poultry (Gavora, 1982), in order to describe laying performance, and to sheep (Groenewald et al., 1995), to predict milk production.

Mathematical models of lactation are also used in genetic evaluation of dairy cattle. The method, referred to as the *test-day model* (TDM), constitutes a part of routine genetic evaluation of herds; Wood's incomplete gamma function, besides other regression models, is used in such analyses. One particular advantage of the method is that the analysis is based on real data, i.e. daily yields actually measured for a cow, *not* those estimated or predicted; this allows considering factors that directly influence daily milk yields (Swalve, 1995).

Besides mathematical functions, artificial intelligence-based models, such as expert systems, genetic algorithms, or artificial neural networks (ANNs), are applied in decision support. ANNs propose an approach that is completely different from that offered by conventional methods, which need an algorithm to be specified and transformed into a computer program. ANNs are capable of parallel data processing, exhibit "associative" access to information, and can process fuzzy, incomplete, or even inconsistent information. They have an ability to adapt, self-organize, and – last, but not least – to learn, which substitutes programming. This specificity of artificial neural networks represents their real power, which can be employed to solve various kinds of problems that individual breeders or corporate farming entities encounter in their day-to-day activity.

Artificial neural networks are mainly used in engineering, economics, or medicine; their application, however, is also imaginable when it comes to food processing (Ni et al., 1994; Paquet et al., 2000; Mei and Chin, 1999) or livestock management (Korthals et al., 1994; Shao et al., 1997; Yang et al., 1999; Suchorski-Tremblay et al., 2001; Kominakis et al., 2002). In dairy cattle farming, for example, artificial neural networks can be successfully used to predict milk production per 305-d lactation (Lacroix et al., 1995), per hectare of pasture, per cow, or per farm (Sanzogni and Kerr, 2001).

Some comparative studies are also worth mentioning, in which the authors tested ANNs predictions against other models (Christy et al., 1995; Heald et al., 2000; Grzesiak et al., 2003). An opinion by Tadeusiewicz (1998) should also be kept in mind that a number of studies, many of which remain unpublished, have not produced a desirable outcome in terms of ANNs application.

So far, ANNs have not been widely used in livestock breeding and management, especially in dairy cattle farming. In this context, it seems reasonable to undertake studies in order to test the method and its usefulness in this field. Monitoring milk yield of a cow provides information on her health and nutritional needs, but also uncovers other specific events; it may also support decisions on culling or selection. Neural networks used in place of sophisticated mathematical formulae may also facilitate analysis of these problems.

The aim of this study was to contrast *methodological* approaches followed in predicting dairy cows lactation yield, i.e. Wood's regression model and artificial neural networks. Within these two approaches we compared: (1) the quality of the neural network against that of Wood's model, (2) predictive capabilities of both models in terms of lactation yield based on daily yields, and (3) predictions by these models against those produced by an official milk recording system.

2. Material and methods

2.1. Datasets

The study was based on daily milk yields automatically recorded on a dairy farm from April 2000 till December 2002. The analyses included data recorded between 5 and 305 days in milk. In all, 137,507 daily yield records have been collected for the first, second, and third lactations of 320 cows. Daily milk recording was performed by means of

Table 1
Datasets used for regressive models and neural network parameter estimation (dataset A) and testing (dataset B); standard deviations in brackets

Lactation	Number of lactations	Number of days in milk	Mean lactation time (days)	Percentage of HF genes	Age at first calving (months)	Average daily milk yield (kg)
Dataset A						
I	173	43,188	257.12 (43.52)	80.41 (23.47)	28.31 (4.93)	18.28 (4.32)
II	166	35,286	238.38 (53.26)	80.24 (25.92)	41.15 (5.27)	22.09 (6.52)
III	129	30,457	251.62 (63.41)	78.12 (29.46)	60.21 (5.88)	22.36 (6.57)
Total	468	108,931	253.51 (48.54)	80.82 (24.55)	41.66 (5.31)	20.66 (5.66)
Dataset B						
I	35	10,041	253.24 (51.33)	77.33 (25.31)	26.63 (4.76)	17.58 (4.16)
II	34	9,894	249.48 (53.26)	76.54 (24.02)	40.01 (5.43)	22.27 (6.86)
III	30	8,641	236.33 (44.82)	73.21 (30.74)	63.55 (6.17)	22.32 (5.90)
Total	99	28,576	248.89 (50.69)	75.81 (26.35)	42.43 (5.39)	20.42 (5.52)

the Alfa-Laval milking equipment installed in the barn and linked to a computer recording data for each cow separately. The cows were milked in the morning and late in the evening. At the intervals of 28–33 days, official milk recording, SYMLEK, SL (ZETO Olsztyn Sp. z o.o., www.zeto.olsztyn.pl) took place on test days; the resulting data were used to determine cumulative lactation yields by calculating average daily yields for two consecutive trials and multiplying the result by days in milk since the previous test day. Thus, the calculated value was added to the previous one, which in consequence produced the cumulative lactation yield. Yields lower than 1 kg of milk per day were ignored. Of these, 108,931 daily yields (dataset A) were used for estimation of regression models parameters and for neural network training, whereas 28,576 records (dataset B) were used for quality testing of the models. Table 1 presents the description of the material.

Division by age and genetic groups was performed using *k*-means cluster analysis with rounded results (Hartigan and Wong, 1979). Eventually, the cows were divided into two genetic subgroups ($\leq$75% HF and 75.1–100% HF), two calving season subgroups (autumn/winter and spring/summer), and two age-at-calving subgroups (18–30 and 31–46 months at first calving, 28–39 and 40–46 months at second calving, and 37–53 and 54–89 months at third calving). Based on this division, eight groups (genotype–age–calving season) were formed for each lactation. Each lactation was also subdivided into 15 periods: (1) 5–20 days in milk, (2) 21–40 days in milk, (3) 41–60 days in milk, ..., (15) 281–305 days in milk.

2.2. Regression models, neural network models, and their quality

We used the function proposed by Wood (1967) to describe the course of lactation:

$$y = at^b \exp(-ct) \tag{1}$$

where y is the average daily milk yield during given time t, t the time expressed in days, a the parameter related to the peak lactation, b the parameter related to the ascending part of the curve between calving and peak lactation, c is the parameter related to the descending part of the curve following the peak lactation. For very low values of c, the value of a is very close to the yield directly after calving. Wood's model assumes the peak lactation yield of $a(b/c)^b \exp(-b)$ appearing in b/c days after calving (Papajcsik and Bodero, 1998). Since a non-linear regression does not guarantee convergence, natural logarithms were taken of both sides of the equation: $\ln(y) = \ln(a) + b\ln(t) - ct + \varepsilon_i$, where ε_i is model random error.

The parameters of Wood's models were estimated with Gauss–Newton method (Hartley, 1961) using the SAS package NLIN procedure of (SAS/STAT, 1989). The parameters were estimated for 24 equations for the first three lactations.

The quality of each regression model was tested analysing the predictions residual distribution for homoscedasticity (constant variance of residues). For this purpose, normality of residual distributions of each model was tested with the Kolmogorov–Smirnov test (with Lilliefors correction) and the residues were tested for autocorrelations using Durbin–Watson statistic and coefficient of autocorrelation (r_a) between the residuals and lagged residuals.

The neural network (NN) model was based on the following predictor variables: x_1 is the HF percentage, x_2 the age at calving in months, x_3 the month of calving (numerically encoded—October: 1, November: 2, December: 3, ..., September: 12), x_4 the days in milk after calving and x_5 is the lactation number. The variables were selected basing on available data, so that the neural network could exploit the same information as the estimated regression models. The actual milk yield on a given day represented the dependent variable (y) in the neural network model.

We used Statistica Neural Network v. 4.0 package (StatSoft, Inc., Tulsa, OK, USA; www.statsoft.com), which enabled us to simulate and analyse a number of networks with various architectures. In all, 25 networks with one or two hidden layers were analysed, from which we chose a perceptron with two hidden layers, 10 and 6 neurons in the first and the second layer, respectively; this particular network exhibited the best quality parameters, mainly $\text{S.D.}_{\text{ratio}}$ (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).

The verification and training datasets were built of data randomly taken from the dataset A, in such a way that the size of the verification dataset (28,576 random records) corresponded to that of the testing dataset, whereas the remaining data represented the training set (80,535 random records).

The testing dataset was used to modify the weights of the network, whilst the verification set was used to control the size of network error during training and, consequently, to control the approximation ability of the network. Input and output data were converted by means of minimax linear conversion function:

$$z_i(t) = \frac{z_i - \min(z_i^n)}{\max(z_i^n) - \min(z_i^n)} \tag{2}$$

where $z_i(t)$ is the observations after conversion, z_i the observations before conversion, $\min(z_i^n)$ and $\max(z_i^n)$ are the dataset minimum and maximum values. Also in the output layer, the results were converted in order to obtain actual values of the trait.

Network training was based on the error back propagation algorithm (Rumelhart et al., 1986). The process went through 40,000 epochs (learning steps), each of which consisted in a single presentation of all cases of the training set and a resulting modification of the network parameters. The learning rate (decreasing from $\eta = 0.9$ to 0.3) and the momentum coefficient (0.5) were adopted for each tested network. We also applied random weight disturbance, which consisted in adding pseudo-random numbers with uniform distribution from the interval ($-n$, n), and input vector disturbance by changing the sequence of example search, in order to reduce the probability of the algorithm falling into a local minimum (Reinhardt and Müller, 1990). The learning process was monitored by means of a graph representing the root mean square (RMS) error for the training and verification datasets as a general network error. A sigmoid function was applied as the activation function.

The following quality measures were adopted for both regression and neural network models:

- R^2_{adj}—adjusted coefficient of determination:

$$R^2_{\text{adj}} = 1 - \frac{\text{MS}_\text{E}}{\text{MS}_\text{T}} \tag{3}$$

  where MS_E is the model error (residual) variance and MS_T is the total model variance.
- RAE—model global relative approximation error:

$$\text{RAE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{\sum_{i=1}^{n} y_i^2}} \tag{4}$$

  where y_i is the actual values and $\hat{y}$ is the regression model predicted values.
- RMS—root mean square error:

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n}} \tag{5}$$

  where n is the number of observations, y_i and $\hat{y}$ are like in the Eq. (4).

Sensitivity analysis of neural network inputs consists of a hands-on evaluation of the loss that occurs when a given predictor is removed. The analysis ranks the variables in order of significance and returns the value of error the network makes working without that given variable. The ratio of the error made without the given variable to the error made with a complete set of variables is also calculated; the value "1" of the ratio shows that the variable is superfluous.

2.3. Predictive capabilities of the models

In order to test the predictive quality of the compared models, records of 99 cows were randomly selected, which represented 28,576 daily yield records for three lactations (dataset B, Table 1). These data were not used to estimate the parameters of either the regression model or the neural network and were collected until the end of the lactation or up to 305 days, if it lasted longer. The neural network predictions were divided into three lactations.

Average daily yields were determined for each stage of lactation, basing on calculations performed for each cow separately, and according to the cow's belonging to a given group. The yields in each group for lactations not longer than 305 days were recalculated for both neural and regression models, summing predictions for subsequent days of lactation.

Next, for each lactation period, a difference was determined between the mean actual daily yields and the mean yields derived from the official milk recording system SYMLEK (SL) and the predictions. Mean absolute and relative differences for entire lactation were summed. Cumulative yields for each 305-d lactation stage (including the entire 305-d lactation yield) were also determined as well as differences between the forecast and the cumulative yield. The predictions by the regression and neural network model were compared with the SL interpolated values, which were derived from official monthly milk recording, and with the mean real (RL) yields.

The following parameters were applied for evaluation of the analysed models prediction quality:

(1) Pearson's coefficient of linear correlation (r) between the predictions by a given model and the real values;
(2) Mean relative prediction error (Ψ) estimated from the formula:

$$\Psi = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}}{y_i}\right| \cdot 100Ȩ \tag{6}$$

where y_i is the actual daily yields or those computed with SYMLEK, $\hat{y}$ the model predicted yields, n is the number of observations.
(3) Theil's inequality coefficient (I^2, Theil, 1979):

$$I^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{\sum_{i=1}^{n} y_i^2} \tag{7}$$

The above coefficient can be broken down into a sum of three coefficients of different error sources:

$$I^2 = I_O^2 + I_B^2 + I_E^2 \tag{8}$$

The components of Eq. (8) are as follows:

I_O^2 represents prediction bias estimated from the formula:

$$I_O^2 = \frac{(\bar{y}_i - \hat{y}_\mathrm{m})^2}{(1/n)\sum_{i=1}^{n} y_i^2}, \tag{9}$$

where $\bar{y}_i$ is the mean actual value and $\hat{y}_m$ is the mean predicted value.
I_B^2 represents error resulting from inadequate flexibility of prediction:

$$I_B^2 = \frac{(\sigma_i - \sigma_\mathrm{p})^2}{(1/n)\sum_{i=1}^{n} y_i^2} \tag{10}$$

where σ_i is the standard deviation of the series of actual values and σ_p is the standard deviation of the series of predicted values.

I_E^2 represents error resulting from insufficient convergence between the directions of prediction changes and the changes in the predicted variable:

$$I_E^2 = \frac{2\sigma_i\sigma_p(1-r)}{(1/n)\sum_{i=1}^{n} y_i^2} \tag{11}$$

where r is the coefficient of linear correlation between given model predictions and actual values; the other symbols are the same as in the previous equations. The quality of Wood's models were compared with the neural networks generated separately for first, second, and third lactation.

3. Results

The value of neural network S.D.$_{\text{ratio}}$ was about 0.43 for the training set and 0.44 for the verification dataset, whereas the coefficient of correlation between the neural network input data and results reached about 0.88 for both the training and the verification sets. Errors of the training and the verification datasets stabilised after about 44,000 epochs of learning.

The regression coefficients estimators presented in Table 2 based on daily milk yields were statistically significant. The coefficients of determination (R^2) for the models ranged between 0.32 in group 8 (first lactation) and 0.79 in group 6 (third lactation). The mean value of this coefficient for Wood's model in subsequent lactations ranged from 0.45 to 0.62, while it was 0.77 for the neural network model (Table 3).

Total relative approximation error (RAE) was 0.031 lower for the network model as compared to Wood's model. Also RMS errors were considerably lower for the network, by about 0.67, and averaged 17.94% for Wood's model and 14.74% for the neural network model (Table 3).

An analysis of residual autocorrelation for individual model has not exhibited significant coefficients of autocorrelation or Durbin–Watson statistic coefficients, except for two groups, 3 and 6, in second lactation. Not all Wood's models have satisfied assumptions for regression models; the model for group 3 did not have normally distributed residuals or equality of variations. The models 4 and 8 exhibited the coefficient of determination below 0.40, and group 4 also did not satisfy the assumption of homoscedasticity of residuals.

The coefficients of correlation between the actual yields and predictions were the highest for the neural network, followed by those obtained for Wood's model and SL (except for second lactation, Table 4). Mean relative prediction errors (Ψ) for the neural network were the lowest, particularly in second lactation, when they were by about 4.42 and 5.22% lower in relation to, respectively, the errors of Wood's model and SL; the errors were quite similar in third lactation. Similar pattern was found for Theil's coefficient. In general, the percentage of each component in the coefficient was similar for the neural network, Wood's model, and SL. The lowest three-lactation average percentage was found for I_O (1.35% for WD and 1.30% for SL, 0.48% for NN), I_B component was slightly higher (12.72% for WD, 13.39% for SL and 8.63% for NN), while the highest percentage was represented by the I_E component (nearly 86% for WD, 85% for SL and 91% for NN).

Table 5 presents average differences between predictions, SL yields, and daily and cumulative yields in subsequent stages of lactation. The predicted yields, after an initial period of underestimation, became higher than the actual daily milk yields. The highest absolute sum of differences in the first lactation was characteristic for SL predictions (18.85 kg), then for WD model (8.87 kg), while for NN was the lowest (5.95 kg). The differences for the SL yields were statistically significant in all stages of first lactations, like for WD model (except for two stages), whereas the differences for NN were significant for less than half of the first lactation stages. In the second lactation, the absolute sum of differences between the actual yields and predictions was the lowest for NN (5.66 kg) and all network predictions exceeded the actual yields in the analysed lactation stages, similarly as WD predictions (absolute sum of differences: 9.55 kg). SL predictions, on the other hand, were overestimated for most lactations stages (absolute sum of differences: 18.25 kg). For the NN model the differences between the actual and predicted yields were, in most stages of second lactation, statistically non-significant (Table 5). In the third lactation, network predictions were mostly overestimated (sum of absolute differences nearly 16 kg), however, they were non-significant for most lactation stages. For the WD models, such overestimation was found in the middle and final stages of lactation (sum of absolute differences: 6.79 kg). SL predictions were overestimated (sum of absolute differences 22.18 kg) and statistically significant for nearly the entire lactation (Table 5).

Table 2
Coefficients of regression (a–e) with standard errors (below) and coefficients of determination (R^2) for analysed models

Group		First lactation				Group	Second lactation				Group	Third lactation			
		a	*b*	*c*	R^2		*a*	*b*	*c*	R^2		*a*	*b*	*c*	R^2
1	**WD**	**12.114**	**0.203**	**0.003**	**0.6016**	1	**20.674**	**0.122**	**0.004**	**0.5307**	1	**21.554**	**0.125**	**0.004**	**0.6484**
$n = 2289$		0.600	0.014	0.000		$n = 4843$	0.728	0.011	0.000		$n = 4150$	0.547	0.008	0.000	
2	**WD**	**16.113**	**0.110**	**0.003**	**0.4412**	2	**21.199**	**0.102**	**0.003**	**0.4722**	2	**20.411**	**0.119**	**0.003**	**0.5416**
$n = 5088$		0.582	0.011	0.000		$n = 4572$	**1.417**	**0.020**	**0.000**		$n = 4842$	1.110	0.016	0.000	
3	**WD**	**15.018**	**0.102**	**0.002**	**0.4159**	3	**16.000**	**0.221**	**0.005**	**0.7405**	3	**19.355**	**0.139**	**0.004**	**0.6743**
$n = 5073$		0.471	0.009	0.000		$n = 3242$	0.549	0.010	0.000		$n = 5153$	0.522	0.008	0.000	
4	**WD**	**12.441**	**0.097**	**0.002**	**0.3586**	4	**19.534**	**0.104**	**0.004**	**0.5146**	4	**14.553**	**0.227**	**0.005**	**0.6419**
$n = 1905$		0.774	0.018	0.000		$n = 3924$	0.776	0.012	0.000		$n = 3399$	2.185	0.044	0.000	
5	**WD**	**16.858**	**0.098**	**0.002**	**0.4215**	5	**21.836**	**0.157**	**0.005**	**0.5561**	5	**20.159**	**0.176**	**0.005**	**0.6015**
$n = 10{,}985$		0.311	0.005	0.000		$n = 5708$	0.660	0.009	0.000		$n = 7719$	0.461	0.007	0.000	
6	**WD**	**16.130**	**0.111**	**0.002**	**0.4397**	6	**20.594**	**0.135**	**0.004**	**0.6984**	6	**20.947**	**0.158**	**0.005**	**0.7869**
$n = 6633$		0.341	0.006	0.000		$n = 3775$	0.476	0.007	0.000		$n = 1240$	0.823	0.012	0.000	
7	**WD**	**16.661**	**0.146**	**0.004**	**0.5181**	7	**21.325**	**0.121**	**0.004**	**0.6081**	7	**19.013**	**0.132**	**0.003**	**0.6284**
$n = 4759$		0.779	0.014	0.000		$n = 5792$	0.572	0.008	0.000		$n = 2305$	0.554	0.009	0.000	
8	**WD**	**17.749**	**0.060**	**0.002**	**0.3161**	8	**22.227**	**0.064**	**0.003**	**0.4719**	8	**17.818**	**0.153**	**0.005**	**0.4697**
$n = 3456$		0.941	0.015	**0.000**		$n = 3430$	1.002	0.014	0.000		$n = 1649$	1.119	0.019	0.000	

Significant differences in bold.

Table 3
Quality parameters Wood's models (WD) and neural network (NN)

Model	RAE	RMS		R^2
		kg	%	
Wood's models				
Lactation I	0.1730	3.2507	17.78	0.45
Lactation II	0.1770	4.08	18.45	0.57
Lactation III	0.1687	3.93	17.58	0.62
I–III	0.1731	3.7090	17.94	0.54
Neural network				
NN	0.14	3.04	14.74	0.77

RAE: global relative approximation error, RMS: root mean square error, R^2: coefficient of determination.

The sum of mean absolute cumulative differences of milk yield was 1750 kg for SL, 830 kg for WD, and 480 kg for NN (Table 5). The 305-d yield, however, did not differ from predictions so much; it was best predicted by an NN model (difference between the actual 305-d and predicted 305-d yields was −96.1 kg), whereas WD-generated difference was −138 kg of milk, and for SL was −313 kg. In the second 305-d lactation, the sum of mean absolute cumulative differences was 2575 kg for SL, 1264 kg for WD, and 712 kg for NN. The total difference between the actual and predicted yields exceeded 2500 kg of milk for SL, 1264 kg for WD, and 712 kg for NN. In the third lactation, the predictions were most accurate; the sum of mean absolute cumulative differences was 1066 kg for SL, 372 kg for WD, and 391 for NN.

The coefficient of correlation between the actual cumulative yields and predictions ranged 0.70–0.97 for the first lactation, 0.87–0.99 for the second lactation, and 0.89–0.99 for the third lactation, depending on the model. Slightly lower correlations were found for SL predictions, while the highest correlations were for NN predictions. Coefficients of correlation between actual 305-d yields and predictions ranged 0.89–0.93 for SL, 0.92–0.94 for WD and 0.94–0.96 for NN, depending on lactation (Table 5). Differences between the coefficients of correlation were statistically non-significant. Table 5 also presents differences between the predictions and real data in cumulated, 100-day periods of lactations. In the first lactation, the highest divergences were recorded for SL, followed by WD and NN. In the second lactation for SL, after initial high differences observable until 200 days (915 and 907 kg milk), the value decreased to −22 kg milk and was lower as compared with WD (757 kg) or NN (449 kg). In the third lactation, the differences for SL ranged from 277 kg for the first 100 days in milk to −360 kg for the last 100 days. Lower variations were observed for WD (from −8.9 to −215 kg of milk) and NN (from −49 to −184 kg).

Table 4
Predictive measures of the analysed models

	r	Mean relative prediction error (Ψ)	Theil's inequality coefficient						
			I_O^2	I_B^2	I_E^2	I^2	$\%I_O^2$	$\%I_B^2$	$\%I_E^2$
Lactation I									
SL	0.6032	17.26	0.0007251	0.0062311	0.0300412	0.0369974	1.96	16.84	81.19
WD	0.6204	16.42	0.0006811	0.0053743	0.0269979	0.0330533	2.06	16.26	81.68
NN	0.8389	13.54	0.0003200	0.0031211	0.0209080	0.0243491	1.31	12.82	85.87
Lactation II									
SL	0.7501	14.67	0.0006272	0.0058427	0.02780041	0.03427031	1.83	17.05	81.12
WD	0.7489	15.47	0.0006721	0.0060443	0.0319903	0.0387061	1.74	15.62	82.65
NN	0.8942	10.25	0.0000092	0.0025593	0.0194741	0.0220422	0.04	11.61	88.35
Lactation III									
SL	0.7756	14.31	0.0000311	0.0018342	0.0274046	0.0292699	0.11	6.27	93.63
WD	0.7923	13.94	0.0000217	0.0013045	0.0232948	0.0246210	0.09	5.30	94.61
NN	0.8617	13.65	0.0000037	0.0000862	0.0241314	0.0242213	0.02	0.36	99.63

I_O^2: prediction bias; I_B^2: prediction flexibility; I_E^2: insufficient convergence between the directions of prediction changes; $I^2 = I_O^2 + I_B^2 + I_E^2$.

Table 5
Average prediction differences of actual daily and cumulative yields and coefficients of correlation for actual and predicted yields by lactation

Days of lactation	Differences between actual, cumulative yields, and prediction						Coefficients of correlation		
	RZ-SL	RZ-WD	RZ-NN	RZ-SL	RZ-WD	RZ-NN	rSL	rWD	rNN
First lactation									
5–20	5.17**	−0.68**	−0.17	69.5	−11.0	−3.2	0.90	0.93	0.95
21–40	−0.63**	0.51**	0.26	92.6	−0.6	2.4	0.91	0.95	0.96
41–60	−0.91**	0.44*	0.22	88.0	8.3	6.7	0.90	0.96	0.97
61–80	−1.31**	−0.07	0.11	107.6	6.4	8.7	0.89	0.92	0.95
81–100	−1.90**	−0.81**	−0.38*	−43.6(314[1])	−9.9(−6,7[1])	0.9(15,6[1])	0.93	0.96	0.92
101–120	−1.05**	−1.06**	−0.68**	14.8	−30.8	−12.7	0.86	0.87	0.92
121–140	−1.00**	−0.89**	−0.56	22.0	−47.4	−22.8	0.87	0.87	0.91
141–160	−0.46*	−0.24	−0.01	−53.2	−52.2	−23.1	0.80	0.93	0.97
161–180	−0.47*	−0.34*	−0.18	−45.8	−59.1	−26.7	0.75	0.93	0.95
181–200	−0.80**	−0.60**	−0.41*	−179.8(−241,9[2])	−71.2(−260,7[2])	−35.0(−120,3[3])	0.70	0.84	0.90
201–220	−0.85**	−0.34*	−0.28	−62.7	−77.7	−40.5	0.87	0.88	0.91
221–240	−1.19**	−0.69**	−0.67**	−221.1	−91.5	−53.8	0.82	0.89	0.89
241–260	−0.77**	−0.68**	−0.64**	−215.3	−104.7	−66.2	0.85	0.86	0.87
261–280	−1.12**	−0.87**	−0.29	−221.3	−121.3	−81.8	0.81	0.85	0.89
281–305	−1.22**	−0.67**	−0.49*	−313.0(−1033,4[3])	−138.0(−533,3[3])	−96.1(−338,3[3])	0.89	0.93	0.94
Total	−8.51	−6.97	−4.17	−961.3	−800.7	−443.1	0.85	0.91	0.93
	18.85	*8.87*	*5.95*	*1750.4*	*830.2*	*480.6*			
Second lactation									
5–20	−0.32	0.15	0.08	99.2	3.5	2.4	0.94	0.98	0.99
21–40	2.45**	1.22**	0.10	199.3	27.8	4.4	0.89	0.90	0.98
41–60	−0.07	0.20	0.03	173.3	32.4	5.5	0.93	0.95	0.98
61–80	0.28	0.49*	0.45	231.6	42.5	14.9	0.91	0.91	0.98
81–100	−1.59**	−0.25	0.03	211.6(915,1[1])	37.9(144,2[1])	15.6(42,9[1])	0.96	0.97	0.98
101–120	−0.36	0.38*	0.44*	115.1	45.6	24.4	0.90	0.91	0.95
121–140	−1.41**	0.63**	0.60**	151.6	58.1	36.2	0.89	0.90	0.89
141–160	−1.60**	0.63**	0.48*	277.3	70.4	45.5	0.87	0.89	0.90
161–180	−1.07**	0.83**	0.40	128.3	87.2	53.8	0.89	0.88	0.97
181–200	−1.29**	0.76*	0.36	234.3(906,7[2])	102.3(363,5[2])	60.9(220,7[2])	0.93	0.90	0.97
201–220	−1.63**	0.68*	0.26	145.5	115.7	66.0	0.89	0.89	0.90
221–240	−1.56**	1.11**	0.63*	109.0	137.8	78.4	0.87	0.89	0.98
241–260	−1.95**	0.65*	0.36	111.0	151.1	85.7	0.87	0.89	0.89
261–280	−1.80**	0.69**	0.71**	−199.2	165.1	100.1	0.89	0.90	0.89
281–305	−0.88**	0.88**	0.73**	−188.4(−22.1[3])	187.0(756.6[3])	118.4(448.7[3])	0.92	0.92	0.94
Total	−12.80	9.05	5.66	1799.7	1264.2	712.3	0.90	0.91	0.95
	18.25	*9.55*	*5.66*	*2574.9*	*1264.2*	*712.3*			

Table 5 (*Continued*)

Days of lactation	Differences between actual, cumulative yields, and prediction						Coefficients of correlation		
	RZ-SL	RZ-WD	RZ-NN	RZ-SL	RZ-WD	RZ-NN	rSL	rWD	rNN
Third lactation									
5–20	−1.46**	−0.51	−0.68	8.4	−7.5	−12.1	0.96	0.97	0.94
21–40	0.08	0.82	0.70	151.6	8.7	−5.4	0.94	0.97	0.98
41–60	0.04	0.00	0.05	87.2	9.1	−2.8	0.95	0.98	0.99
61–80	−2.41**	−0.44	0.12	18.9	0.2	−4.6	0.92	0.98	0.98
81–100	−1.95**	−1.00	−1.45	11.0(277.0[1])	−19.4(−8.9[1])	−16.4(−41.3[1])	0.96	0.99	0.99
101–120	−2.62**	−0.67	−1.23	128.4	−32.4	−23.7	0.96	0.98	0.98
121–140	−3.07**	−0.74	−0.35	−49.2	−47.2	−36.3	0.94	0.96	0.97
141–160	−3.45**	−0.08	−0.58	10.8	−48.7	−39.4	0.90	0.90	0.93
161–180	−1.13	0.22	−0.05	36.3	−44.1	−39.4	0.96	0.96	0.97
181–200	−0.43	0.06	0.18	−104.5(21.8[2])	−42.9(−215.4[2])	−45.2(−184.0[2])	0.95	0.95	0.95
201–220	−1.34	0.24	−0.22	−89.5	−38.2	−48.2	0.96	0.97	0.99
221–240	−0.69	0.78	0.12	50.3	−22.4	−38.8	0.94	0.95	0.95
241–260	0.80	0.58	−0.04	−120.4	−10.9	−29.3	0.89	0.93	0.93
261–280	−1.55	−0.08	−0.09	−87.6	−12.3	−25.9	0.92	0.93	0.90
281–305	−1.17	−0.58	−0.25	−112.4(−359.6[3])	−27.7(−111.5[3])	−23.2(−165.4[3])	0.93	0.94	0.96
Total	−20.34	−1.40	−3.77	−60.8	−335.8	−390.7	0.94	0.96	0.96
	22.18	*6.79*	*15.94*	*1066.4*	*371.9*	*390.7*			

Absolute sums in italics. Differences for cumulative yields predictions in periods: (superscript 1) 5–100 days, (superscript 2) 101–200 days and (subscript 3) 201–305 days.

* $P<0.05$.

** $P<0.01$.

Table 6
Sensitivity analysis for neural network

Variables	Days	Hf	Age	Month	Lactation
Range	1	5	4	2	3
Error	5.59	4.2	4.34	5.03	4.42
Ratio	1.41	1.06	1.1	1.27	1.12

Range: range of variables according to validity; error: indicate size of loss after rejection of given variable; ratio: relation of error to total error received when use all variables; days: days of milk; HF: percentage of hf gene; age: age of calving; month: month of calving.

Sensitivity analysis for the neural network model revealed that days in milk (days, Table 6) was the most significant variable that influenced daily milk yield predictions. Removing this variable would lead to a 1.41-fold increase in the accumulated network error (ratio). The second important variable was month of calving (MONTH, 1.27-fold accumulated network error increase). Removing HF percentage variable would lead to the least increase in this error (1.06 times).

4. Discussion

Application of a neural network was definitely an easy task, since the data could be directly presented to the model. In order to estimate the parameters of Wood's lactation curve, it was necessary to divide the material into age groups, genetic groups, calving-season groups, and lactation groups, which in consequence produced 24 equations. A more detailed division of the material, i.e. distinguishing all ages (in months), more detailed breakdown of genetic groups, or substitution of calving season with the month of calving would require much larger material, so that estimation of the large number of regression equations could at all take place. This, for single farms, is virtually impossible and the neural network is an optimum solution for such situations, which has been demonstrated by the presented quality measures and milk yield predictions.

4.1. Quality evaluation of the regression model and ANN

The decision as to which network model would be best for predictions was based on the coefficient of correlation between the input and output data and the ratio of standard deviations of errors and input data ($\text{S.D.}_{\text{ratio}}$). The values of $\text{S.D.}_{\text{ratio}}$ (less than 0.4) indicate a good quality of the selected network; for $\text{S.D.}_{\text{ratio}}$ higher than 0.7, a network would not be useful for predicting, whereas the index lower than 0.1 means that the network would be nearly ideal (Statistica Neural Network, 1998).

The training we applied to the network can be considered as sufficient, which was demonstrated by the general network error whose converted value reached about 0.1. Dayhoff (1990) stated that the general network error lower than 0.1 means that the network has been well trained. Skapura (1996) suggested that even an error lower than 0.2 is enough to assume sufficient training.

Optimising the process of training also requires that all or some network parameters should be tuned. According to some authors (Salehi et al., 1998), the most important parameters are the number of learning epochs and the type of data presentation. Those of lesser importance include learning rate, whilst the momentum coefficient is usually unimportant. According to Stein (1993), the process of training is also shaped by data preprocessing, especially for nominal or rank variables. We have applied the initial procedure, *minimax*, which allowed the network to receive proportional stimuli (milk yield, lactation number). The calving season variable was presented to the networks in the form of subsequent months; the interval scale was converted into a rank scale basing on the fact that cows calving during autumn/winter had a higher yield; hence the autumn/winter months of calving were ranked as 1–6 and spring/summer months as 7–12.

For a proper learning process, an appropriate representation of data relevant for the given problem is necessary, which has been raised by Lacroix et al. (1997). Appropriate representation guarantees that the network-generated prognoses will be relevant for the given population. Using additional information on the population of cows may contribute to a better prediction produced by the network. In our study, we have intentionally ignored any additional information, so that the predictor variables would be comparable to those used in the regression model. The fact that yields of different cows on the same days of lactation were similar posed a problem in arriving at a satisfactory network.

The single output variable applied in our model has contributed to a better network optimisation, particularly as large datasets were involved, contrary to a network with multiple outputs (Lacroix et al., 1995).

The coefficients of determination (R^2) for the WD models showed their fitness being at the edge of applicability (0.45, 0.57, and 0.62 for respective lactations I–III). The R^2 coefficients for the network (Table 2) were higher and similar to those for models based on monthly yields, for example, Wood (1967) obtained $R^2 = 0.79$, and Olori et al. (1999) even $R^2 = 0.94$. On the other hand, Freeze and Richards (1992) recorded $R^2 = 0.51$ for first lactation. Olori et al. (1999) also stress that R^2 higher than 0.70 indicates a goodness of fitness, whereas values lower than 0.4 disqualify the model. Some authors (Elston et al., 1989; Perochon et al., 1996) suggest that individual daily yield changes and lactation peak are important for the model fitness. Goodness of fit of models slightly increases with the number of parameters (Jamrozik et al., 1997).

In our study, we have also estimated relative approximation error (RAE) and root mean square error (RMS), which have shown a better quality of the network compared to linear models. Farhangfar et al. (2000) achieved RMS = 3.63 for Wood's model of daily cows yield, which was about 0.6 higher than the RMS we have found for the network and slightly lower than that for our WD models. Salehi et al. (1998) estimated RMS (expressed as percentage of population average) ranging 6–7 to 12%, or even 21% depending on the network and herd milk production. Our estimations of RMS generally remained at a slightly higher level (approx. 15%); however, these values are difficult to compare, as the network described by Salehi et al. (1998) was applied to 305-d milk yield prediction and was built on a much larger set of input variables.

4.2. Prognostic evaluation of the models

Better quality parameters of the network (coefficient of correlation, mean relative prediction error, and Theil's inequality coefficient) also meant more accurate predictions, as compared with a regression model (Table 5). What is remarkable, the highest percentage of Theil's coefficient is represented by I_E, an error resulting from a lack of full convergence in the direction of changes between actual and predicted values, particularly for the neural network. The network overestimated yields in some periods of lactation while underestimating them in the other periods. As far as the WD models and SL are concerned, their certain insufficient flexibility of prediction can be seen (I_B), or an insufficient degree that the model should project the variation of the output variable (Theil, 1979). The variability of predictions was different from the actual changes in the predicted variable.

Especially during the initial stages of lactations, network yield predictions were more accurate that those by WD (Table 5). Bearing in mind that the milk yields recorded in the initial part of lactation may represent a good basis to forecast the yield of the remaining period of lactation (Żuk et al., 1981), such behaviour of the network is highly desirable. Lacroix et al. (1995) observed that predictive abilities of ANNs either improved beginning from the first until the last lactation stage or remained more accurate only during the initial stages of lactation, depending on the network type. Rowlands et al. (1982) found that WD-based predictions of daily yields until 84 days of lactation were underestimated, whereas in the later period (85–175 days) the model produced overpredictions, like our Wood's model. Cobby and Le Du (1978), also using Wood's model, obtained slightly overestimated predictions in the initial period of lactation (until 70 days), whereas underestimated predictions were found between 71 and 140 days of lactation. This divergence may have resulted from the calving season effect, different lactation persistence, or from the differences in the time of reaching peak lactation.

Sensitivity analysis of the ANN model allowed us to stress the fundamental importance of the variable "days", which denotes a subsequent day in milk (Table 6). An effect of days in milk is treated in linear models as a fixed effect; it can be also used, however, as a covariate, which consequently does not lead to divergences in trait variation description (Strabel and Szwaczkowski, 1995). Somewhat surprising may be the fact that "calving age" has a lower importance, although many authors distinguish this factor for its effect on milk yield, particularly in the first lactation. Hence, some authors divide cows into many age groups, e.g., more than 10 (Pirlo et al., 2000), about 5 (Ptak and Schaeffer, 1993), or even groups are formed by age in months (Bormann et al., 2002). Some authors establish fewer age groups, e.g., two (Tsuruta et al., 2002; Koonawootrittriron et al., 2001). In our study, forming fewer age groups resulted from the size and character of the datasets. HF percentage has not significantly influenced daily yield predictions.

The lactation yield estimations based on the official milk recording system (SYMLEK) were worse that the predictions by the ANN or Wood's model alike; however, 305-d yields did not differ so much. The coefficients of correlation between predicted and actual yields indicate that predictions fit well to reality ($r > 0.90$). Wood (1967) obtained 0.95

correlation for predicted and test yields. Ali and Schaeffer (1987) reported higher correlations ($r = 0.98$) for 305-d yield of first lactation, whereas Farhangfar et al. (2000)—even higher coefficient of correlation (0.99) for Wood's model. Other authors, who studied other models, reported various levels of correlation, e.g., Badner and Anderson (1985): $r = 0.10$; Nelder (1966): $r = 0.38$; Jenkins and Ferrel (1984): $r = 0.72$; Guo and Swalve (1995): $r = 0.99$.

The study by Sanzogni and Kerr (2001) may serve as a confirmation of the better quality of ANNs as compared with multiple regression. The authors achieved a better fit of their network models ($R^2 = 0.80$–0.91) than that of a multiple regression models ($R^2 = 0.78$–0.88). Grzesiak et al. (2003) obtained a mean ANN 305-d yield prediction that was only 13 kg lower than the actual average for the analysed cows; whereas the multiple regression prediction was, on average, worse by 91 kg of milk.

There are some difficulties in the application of daily yields in the herd management and genetic evaluation. More and more dairy farms in many countries are implementing daily yield recording. The yields can be next processed in centralised milk recording agencies, so a theoretical possibility exists to utilise the data for generating the lactation curve (Swalve, 1995). The problem, however, arises from the size of such datasets, particularly in terms of estimations of model parameters (Jamrozik et al., 1997).

A neural network can be modified (trained) basing on new data, whereas new data fed to a regression model basically needs a repeated estimation process. ANN can be prepared using fewer observations and, as has been demonstrated in this study, such a model will offer better quality than a regression alternative. Also, ANN superiority results from the fact that there is no need to distinguish genetic or age groups, or barn production levels, since such variables can be fed to the network as continuous predictors. When estimating regression model parameters, one needs to divide cows into specific groups, each of which needs a different equation. In our study, 24 regression equations have been formed, using only limited division into groups. An ANN can use additional information, e.g., somatic cell count, diseases and disorders, and other measurable factors; however, such data have not been included in this study, so that to unify the environments of the models for the sake of their comparability.

In the case of ANN, it is not required to meet the assumptions needed in regression models estimation (normal distributions of residuals, lack of autocorrelation of residuals, or homoscedasticity of residuals), thus making application of ANNs easier.

5. Conclusions

The study allows concluding that our neural network model has met the expectations in terms of its predictive abilities; its application in lactation modelling and 305-d yield prediction seems justified due to the following:

(1) Quality parameters of the ANN were better compared with those estimated for Wood's model and SYMLEK.
(2) Prognostic parameters of the ANN have revealed a lower prediction error.
(3) The ANN-generated milk yield forecasts were more accurate than those by Wood's model and SYMLEK.
(4) The use of an ANN did not require meeting the assumptions of regression models.
(5) Both generating and using the network model was easier than regression models.

It should be stressed that the predictions generated both by the neural network and by the regression models were better that those estimated with the official milk recording system (SYMLEK).

References

Ali, T.E., Schaeffer, L.R., 1987. Accounting for covariances among test day milk yields in dairy cows. Can. J. Anim. Sci. 67, 637–644.

Badner, G.B., Anderson, C.R., 1985. Evaluation of five lactation curve models fitted from daily milk weights. J. Dairy Sci. 68 (1), 226 (Abstract).

Bormann, J., Wiggans, G.R., Druet, T., Gengler, N., 2002. Estimating effects of permanent environment, lactation stage, age, and pregnancy on test-day yield. J. Dairy Sci. 85 (1), 21–42.

Christy, P.S., Tervonen, O., Scheithauer, W.B., Forbes, G.S., 1995. Use of a neural network and multiple regression model to predict histologic grade of astrocytoma from MRI appearances. Neuroradiology 37 (2), 89–93.

Cobby, J.M., Le Du, Y.L.P., 1978. On fitting curves to lactation data. Anim. Prod. 26 (2), 127–133.

Dayhoff, J.E., 1990. Neural Network Architectures: An introduction. Van Nostrand Reinhold, New York, NY.

Dekkers, J.C.M., Ten Hag, J.H., Weersink, A., 1998. Economic aspect of persistency of lactation in dairy cattle. Livest. Prod. Sci. 57, 237–252.

Elston, D.A., Glasbey, C.A., Nelson, D.R., 1989. Non-parametric lactation curves. Anim. Prod. 48, 331–339.

Farhangfar, H., Rowlinson, P., Willis, M.B., 2000. Estimation of lactation curve parameters for Iranian Holstein dairy cows using non-linear models. http://www.bsas.org.uk/meetings/annlproc/PDF2000/109.pdf.

Freeze, B.S., Richards, T.J., 1992. Lactation curve estimation for use in economic optimization models in the dairy industry. J. Dairy Sci. 75, 2984–2989.

Gavora, J.E., 1982. Comparison of three mathematical model of egg production. J. Poult. Sci. 23, 339–348.

Groenewald, P.C.N., Ferreira, A.V., Van der Merwe, H.J., Slippers, S.C., 1995. A mathematical model for describing and predicting the lactation curve of Merino ewes. Anim. Sci. 61, 95–101.

Grzesiak, W., Lacroix, R., Wójcik, J., Błaszczyk, P., 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. Can. J. Anim. Sci. 83 (2), 307–310.

Guo, Z., Swalve, H.H., 1995. Modelling of the lactation curve as a sub-model in the evaluation of test day records. Interbull. Bull. 11, 22–25.

Hartigan, J.A., Wong, M.A., 1979. Algorithm 136. A *k*-means clustering algorithm. Appl. Stat. 28, 100.

Hartley, H.O., 1961. The modified Gauss–Newton method for the fitting of nonlinear regression functions by least squares. Technometrics 3, 269–280.

Heald, C.W., Kim, T., Sischo, W.M., Cooper, J.B., Wolfgang, D.R., 2000. A computerized mastitis decision aid using farm-based records: an artificial neural network approach. J. Dairy Sci. 83 (4), 711–720.

Jakobsen, J., Madsen, P., Jensen, J., Peersen, J., Cristiensen, L.G., Sorensen, D.A., 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. J. Dairy Sci. 85, 1607–1616.

Jamrozik, J., Kistemaker, G.J., Dekkers, J.C.M., Schaeffer, L.R., 1997. Comparison of possible covariates for use in a random regression model for analyses of test day yields. J. Dairy Sci. 80, 2550–2556.

Jenkins, T.G., Ferrel, C.L., 1984. A note on lactation curves of crossbred cows. Anim. Prod. 39, 479.

Kellogg, D.W., Urquhart, N.S., Ortega, A.J., 1977. Estimating Holstein lactation curves with a gamma curve. J. Dairy Sci. 60, 1308–1315.

Kominakis, A.P., Abas, Z., Maltaris, I., Rogdakis, E., 2002. A preliminary study of the application of artificial neural to prediction of milk yield in dairy sheep. Comp. Electron. Agric. 35, 35–48.

Koonawootrittriron, S., Elzo, M.A., Tumwasorn, S., Sintala, W., 2001. Lactation curves and prediction of daily and accumulated milk yields in a multibreed dairy herd in Thailand using all daily records. Thai. J. Agric. Sci. 34, 123–139.

Korthals, R.L., Hahn, G.L., Nienaber, J.A., 1994. Evaluation of neural networks as a tool for management of swine environments. Trans. ASAE 37 (4), 1295–1299.

Lacroix, R., Wade, K.M., Kok, R., Hayes, J.F., 1995. Prediction of cow performance with a connectionist model. Trans. ASAE 38 (5), 1573–1579.

Lacroix, R., Salehi, F., Yang, X.Z., Wade, K.M., 1997. Effect of data preprocessing on the performance of artificial neural networks for dairy yield prediction and cow culling classification. Trans. ASAE 40 (3), 839–846.

Leon-Velarde, C.U., McMillan, I., Gentry, R.D., Wilton, W., 1995. Models for estimating typical lactation curves in dairy cattle. J. Anim. Breed. Genet. 112, 333–340.

Mei, J.S., Chin, B.C., 1999. Experimental studies of network parameters and operational variables on recurrent backpropagation neural network adaptive control of penicillin acylase fermentation by *Arthrobacter viscosus*. Bioprocess Eng. 21, 69–76.

Ni, H., Gunasekaram, S., Bogenrief, D., Olson, N.F., 1994. Predicting cheese quality parameters using neural networks. ASAE paper no. 94-3560. ASAE, St. Joseph, Michigan.

Nelder, J.A., 1966. Inverse polynomials, a useful group of multi-factor response functions. Biometrics 22, 128–141.

Olori, V.E., Brotherstone, S., Hill, W.G., McGuirk, B.J., 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. Livest. Prod. Sci. 58, 55–63.

Papajcsik, I.A., Bodero, J., 1998. Modelling lactation curves of Friesian Cows in a subtropical climate. Anim. Prod. 47, 201–207.

Paquet, J., Lacroix, C., Thibault, J., 2000. Modeling of pH and acidity for industrial cheese production. J. Dairy Sci. 83 (11), 2393–2409.

Perochon, L., Coulon, J.B., Lescourret, F., 1996. Modelling lactation curves of dairy cattle with emphasis on individual variability. Anim. Sci. 63, 189–200.

Pirlo, G., Miglior, F., Speroni, M., 2000. Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. J. Dairy Sci. 83, 603–608.

Ptak, E., Schaeffer, L.R., 1993. Use of test day yields for genetic evaluation of dairy sire and cows. Livest. Prod. Sci. 34, 23–34.

Reinhardt, J., Müller, B., 1990. Neural Networks. An Introduction. Springer-Verlag, Berlin, pp. 132–137.

Rowlands, G.J., Lucey, S., Russell, A.M., 1982. A comparison of different models of the lactation curve in dairy cattle. Anim. Prod. 35, 135–144.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536.

Salehi, F., Lacroix, R., Wade, K.M., 1998. Effects of learning parameters and data presentation on the performance of backpropagation networks for milk yield prediction. Trans. ASAE 41 (1), 253–259.

Sanzogni, L., Kerr, D.V., 2001. Milk production estimates using feed forward artificial neural networks. Comp. Electron. Agric. 32 (1), 21–30.

SAS/STAT, 1989. SAS/STAT User's Guide. Version 6, fourth ed., vol. 1. SAS Institute, Inc., Cary, NC.

Schaeffer, L.R., Jamrozik, J., 1996. Multiple trait prediction of lactation yields for dairy cows. J. Dairy Sci. 79, 2044–2055.

Shanks, R.D., Berger, P.J., Freeman, A.E., Dickinson, F.N., 1981. Genetics aspect of lactation curves. J. Dairy Sci. 64, 1852–1860.

Shao, J., Xin, H., Harmon, J.D., 1997. Neural network analysis of postural behavior of young swine to determine the IR thermal comfort state. Trans. ASAE 40 (3), 755–760.

Skapura, D.M., 1996. Building neural networks. ACM Press, Addison-Wesley Publishing Co., New York, NY.

Sölkner, J., Fuchs, W., 1987. A comparison of different measures of persistency with special respect to variation of test day yields. Livest. Prod. Sci. 16, 305–319.

Statistica Neural Network, 1998. User Manual. Quik Reference. StatSoft Inc., Tulsa, OK.

Stein, R., 1993. Selecting data for neural network. AI Expert 8 (2), 42–47.

Strabel, T., Szwaczkowski, T., 1995. Certain nongenetic effects on test day milk yields. Anim. Sci. Pap. Rep. 13 (2/3), 55–64.

Suchorski-Tremblay, A.M., Kok, R., Thompson, J.J., 2001. Modeling horse hoof cracking with artificial neural networks. Can. Biosyst. Eng. 43, 7.15–7.22.

Swalve, H.H., 1995. Test day models in the analysis of dairy production data. Arch. Tierz. 38, 591–612.

Swalve, H.H., 2000. Theoretical basis and computational methods for different test-day genetic evaluation methods. J. Dairy Sci. 83, 1115–1124.

Tadeusiewicz, R., 1998. Elementary Introduction to Neural Networks with Example Programmes. AOW PLJ, Warszawa (In Polish).

Theil, H., 1979. World income inequality. Econ. Lett. 2, 99–102.

Tsuruta, S., Misztal, I., Klei, L., Lawlor, T.J., 2002. Analysis of age-specific predicted transmitting abilities for final scores in Holsteins with a random regression model. J. Dairy Sci. 85, 1324–1330.

Wood, P.D.P., 1967. Algebraic model of the lactation curve in cattle. Nature 216, 164–165.

Yang, X.Z., Lacroix, R., Wade, K.M., 1999. Neural detection of mastitis from dairy herd improvement records. Trans. ASAE 42 (4), 1063–1071.

Żuk, B., Filistowicz, A., Szyszkowski, L., Hibner, A., 1981. Corrections for predicting yields of milk, fat, and protein for 305-day lactation basing on short-term yields of Red-and-White and Black-and-White cows in south-western Poland. Rocz. Nauk Rol. Ser. B 101 (3), 87–107 (In Polish).