



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## Prediction of Second Parity Milk Yield and Fat Percentage of Dairy Cows Based on First Parity Information Using Neural Network System

<sup>1</sup>P. Hosseinia, <sup>2</sup>M. Edrisi, <sup>1</sup>M.A. Edriss and <sup>3</sup>M.A. Nilforooshan

<sup>1</sup>Department of Animal Science, Isfahan University of Technology, Isfahan, Iran

<sup>2</sup>Departments of IT and Biomedical Engineering, University of Isfahan, Isfahan, Iran

<sup>3</sup>Department of Animal Breeding and Genetics,  
Swedish University of Agricultural Sciences, Uppsala, Sweden

---

**Abstract:** Neural network system can be used as a decision making support system in dairy industry as well as other industries. It can help breeders to predict future yield of dairy cows based on uncorrelated and orthogonalized available information and making selection decisions. Data from 4 medium to large sized dairy farms in Isfahan, Iran, were used. From 1880 available records of first and second parities, 1850 records were used for training a back propagation artificial neural network system and 30 randomly chosen records (not used in the system training step) were introduced to the trained neural network system for its evaluation. The results of the simulation showed that there was no significant difference between the observed and the predicted second parity milk yield and fat percentage ( $p>0.05$ ). The major use of this predictive process is to make accurate selection decisions which are based on prior knowledge of the outcomes.

**Key words:** Neural network, prediction, correlation, milk yield, fat percentage

---

### INTRODUCTION

Artificial Neural Network (ANN), the same as biological neural network, is made up of a set of neurons. These neurons process the presented input and matching output in a supervised manner and make extract non-linear relationships between those input and output. Information processing in ANN is a form of parallel. Actually, ANN is a form of simulated human central nervous system (Wildberger, 1990; Adamczyk *et al.*, 2005). ANN offers an approach that is completely different from those offered by previous conventional methods, which need an algorithm to be specified and transformed by a computer program (Grzesiak *et al.*, 2006). Although linear algebraic methods have wide ranges of application in agriculture, they have some inherent restrictions. Identifying patterns and relationships between input and corresponding output in a sample data depends on the facts that the optimal net performance is dependent to recognition and extraction of non-linear relationships through training an ANN structure (Lacroix *et al.*, 1995). Furthermore, there are some common methods for improving the net performance such as: finding optimum net structure, sufficient number of training cycles and using various kinds of inputs to adjust and evaluate

learning parameters (Lacroix *et al.*, 1997). Network learning ability is more depended to its presented learning examples and patterns (Dayhoff, 1990). However, divergence in performance was often reported in array agriculture, due to over modification of the net structure (Sablani *et al.*, 1995; Yang *et al.*, 1999). Preparing neural network to perform can occur in several stages. In learning stage, ANN can acquire the ability of finding and learning hidden non-linear relationships between input and corresponding target variables, by introducing the presented data to the network as its training data set, through its supervised learning process (Lacroix *et al.*, 1997). Using these relationships in the simulation stage, ANN can anticipate the output by a complex biological system from known discrete/continuous inputs. Thus, for approaching to an optimal learning, a good data presentation is important. Among various kinds of ANN, back propagation ANN is in use in different areas (Lacroix *et al.*, 1997). ANN has been used for estimation of aflatoxin contamination in pre-harvest peanuts (Parmer *et al.*, 1997), plant growth (Brons *et al.*, 1993), prediction of the response of corn yield to nitrogen stress (Kwonnoh *et al.*, 2004), calcium concentration and rheological characteristics of Mozzarella cheese (Ganesan *et al.*, 2004), classification of heart abnormalities

(Md Saad *et al.*, 2007) and flowering and physiological maturity of soybean (Elizondo *et al.*, 1994). It was also employed for detection of susceptible animals to mastitis (López-Benavide *et al.*, 2003; Yang *et al.*, 1999), estimating meat quality of beef (Brethour, 1994), prediction of slaughter value for bulls (Adamczyk *et al.*, 2005), evaluation of physiological status of cows (Molenda *et al.*, 2001), predicting genetic value of domestic animals for herd management practices (Korthals *et al.*, 1994) and especially for evaluation and classification of dairy cows for 305 day production in one lactation (Lacroix *et al.*, 1995; Salehi *et al.*, 1998).

Milk production of dairy cattle is affected by linear and non-linear interactions between cows' genetic merit and the environment (Kominakis *et al.*, 2002). In dairy industry, forecasting bulls' genetic merit in milk traits based on daughters' production is important for identifying and selecting superior bulls. Consequently, as a result of faster evaluation of bulls, their semen can be proved in a shorter period of time (Salehi *et al.*, 2000). However, in special cases such as in dairy industry where the generation interval is about 8 years, rapid detection of superior animals leads to a shorter generation interval and will speed up genetic progress (Lacroix *et al.*, 1995). Accuracy rate in identifying superior cows has a special role, because overestimating dairy cows may lead to extra costs in maintenance, feeding and hygiene for cows of lower genetic values. While, ANN has been applied for detection of mastitis (Yang *et al.*, 1999) and prediction of full 305 day yield from test-day and partial lactation records successfully (Grzesiak *et al.*, 2006; Lacroix *et al.*, 1995; Salehi *et al.*, 1998), the aim of this study was to predict second parity 305 day performance by using available information on the first parity.

Most of dairy farmers are interested in a way for anticipating the yield of their cows in the future. By the implementation of this method in dairy industry, dairy farmers can find from the first parity information that how their cows will produce in the next lactation. They can select more producer cows and because a proportion of this superiority will be transferred to the next generation, it would lead to herd genetic improvement.

The objective of this study was to investigate the ability and accuracy of ANN in assessing and predicting second parity milk yield and fat percentage of dairy cows adjusted for 305 days in milk, from first parity information, for selecting high producing cows as prospective producers and the parents of the next generation.

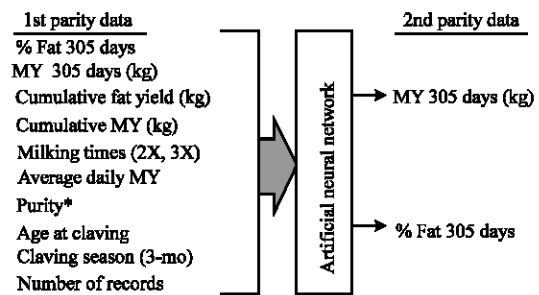
**MATERIALS AND METHODS**

The data was provided by the Animal Husbandry Division, Agricultural Organization of the Ministry of

Agriculture in Isfahan, Iran, which consisted of the collected information from 32 herds during 1995 to 2002. From the available herds, four medium to large sized Holstein herds were selected randomly for final investigations. Records were restricted to those cows which had complete pedigree and production records for both first and second parities. Followed by this restriction, a sample of 1880 cows with records was made available for further studies. The sample data was consisted of cows' registration number, purity (% Holstein blood which was 65.5±19.43 in the sample), first and second parities milk yield and fat percentage, corrected for 305 days in milk and some other information on the first parity of the cows. Then data structure rechecked and the data was introduced to MATLAB (2006) software for further process.

In order to train ANN, ten variables related to the first parity and two variables of the second parity corresponding to the first parity of the individual cow were introduced to the system (MATLAB, 2006) as input and output variables, respectively. Then the minimum and maximum values of each variable were mapped to the mean and standard deviation of 0 and 1, respectively. The related input-output formation is shown in Fig. 1.

Through 1880 records, 1850 records were used for ANN training and 30 records were selected randomly for testing the simulated system (ANN). To study the effect of assorted and unsorted data on ANN training, two ANNs were simulated; ANN1 for which 1850 training records were assorted and categorized into nine production levels (Table 1) and ANN2 for which records were unsorted. The whole data were divided into four subsets: 925 records as a training set, 500 records as an evaluation set, 425 record as testing set and 30 records for the simulation set. Then these data sets were introduced to ANN as a matrix in which any column had a mean and standard deviation of 0 and 1, respectively.



\* % Holstein blood; MY: Milk Yield

Fig. 1: Variables used in the experimental data sets

In order to construct the network, the neural network toolbox was used (MATLAB, 2006). The constructed network was a back propagation Artificial Neural Network (ANN) with three layers of input, hidden and output. The layers had 10, 10 and 2 neurons, respectively. Figure 2 shows how neurons in different layers of an ANN connect together. The tangent hyperbolic transfer function was applied for input and hidden layers and the purline transfer function was applied for the output layer (MATLAB, 2006). The net learning function updated weight and bias values conformable to Levenberd-Marquerdt optimization algorithm (Hagan and Menhaj, 1994). The net was trained in 1E+5 cycles of element processing which were including epoch = 6, goal = 1E-10, MSE = 0.54 and gradient = 558.59.

Where; Epoch: A single pass through the sequence of all input vectors; Goal: Performance is minimized to the goal parameter; MSE: Mean square error of performances; Gradient: A back propagation algorithm which adjusts weights in the steepest descent direction (negative of the gradients).

T-test and correlations were employed for comparing ANN1 and ANN2 predictions with observed values. Also, ANN1 and ANN2 efficiencies were compared by  $\phi$  value, described by Kominakis *et al.* (2002).

$$\phi = (\delta (oy - py)/\text{mean} (oy)) \times 100$$

Where,  $\phi$  is the ratio between the standard deviation ( $\delta$ ) of the differences between the observed (oy) and predicted yields (py) divided by the mean of observed values multiplied by 100.

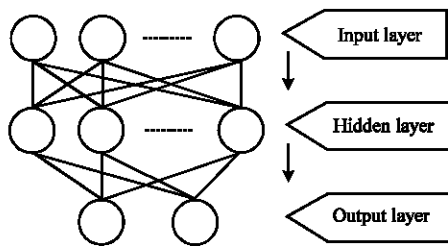


Fig. 2: Neuron connections in a systematic ANN

Table 1: Classification of milk yield for ANN1 training

Milk yield (kg)	Level
<2000	1
2000-3000	2
3000-4000	3
4000-5000	4
5000-6000	5
6000-7000	6
7000-8000	7
8000-9000	8
>9000	9

ANN1: Artificial neural network prediction for the assorted data

## RESULTS AND DISCUSSION

Table 2 results show that there was no significant difference between both ANN1 and ANN2 predictions and the observed values ( $p>0.05$ ), which shows that the results of both ANNs are reliable for both milk yield and fat percent traits. The high correlations showed that the predicted average for milk yield and fat percentage were close to the observed values (Table 3). This shows that ANNs are reliable as a decision support system that helps breeders to choose a cow to be left or culled from herd.

Lower  $\phi$  values and higher correlations ( $r$ ) for ANN1 showed that it could have a higher accuracy of prediction relative to ANN2. Modification of learning or training parameters and the method of data presentation can considerably influence the network performance (Salehi *et al.*, 1998). The performance of ANN1 with assorted data seems to be more justified and as it has been stated by Kominakis *et al.* (2002), prediction accuracy increases by assorted training data relative to the use of unsorted data. The assorted data may cause a better trend, proper updated weights and less bias in the network system and as a result, to high correlation coefficients with the observed values. This may be due to the fact that as good as the presented data to ANN is well-structured, there would be a better training for ANN. i.e., orthogonizing and classification of input vectors, provides better predictions and improves the predictive ability of ANNs due to the fact that they perform particularly well in interpolation (Lacroix *et al.*, 1995).

Table 2: Data structure for the observed and predicted (ANN1 and ANN2) data

Trait	Data	Min	Max	Mean	SD	CV
Milk yield (kg)	OBS	4646.80	9790.80	7342.00	1348.00	18.35
	ANN1	6305.40	9139.80	7400.00	784.35	10.59
	ANN2	6071.57	9281.47	7428.00	897.53	12.08
Fat (%)	OBS	2.01	4.31	2.67	0.55	20.84
	ANN1	2.40	3.51	2.74	0.27	9.88
	ANN2	2.39	3.38	2.68	0.24	9.15

OBS: Observed value; ANN1: Artificial neural network prediction for the assorted data; ANN2: Artificial neural network prediction for the unsorted data

Table 3: Correlations and comparisons between the observed and predicted (ANN1 and ANN2) data

Trait	Data	OBS		
		t	$\phi$	r
Milk yield (kg)	ANN1	0.20 <sup>ns</sup>	11.10	0.82 <sup>***</sup>
	ANN2	0.49 <sup>ns</sup>	11.90	0.77 <sup>***</sup>
Fat (%)	ANN1	0.59 <sup>ns</sup>	12.56	0.90 <sup>***</sup>
	ANN2	0.11 <sup>ns</sup>	22.55	0.85 <sup>***</sup>

\*\*\*  $p<0.001$ ; ns:  $p>0.05$ ; OBS: Observed value; ANN1: Artificial neural network prediction for the assorted data; ANN2: Artificial neural network prediction for the unsorted data; t: t-value for the mean difference between the observed (oy) and predicted (py) data;  $\phi = (\delta (oy-py)/\text{mean} (oy)) \times 100$ ; r: Correlation coefficient

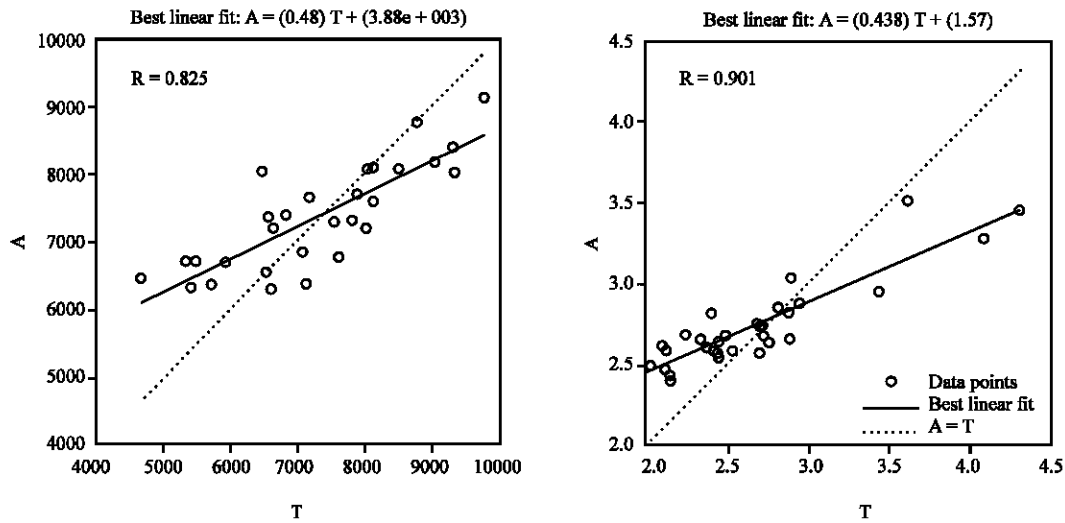


Fig. 3: Regressions of ANN1 predicted on the observed data for second parity milk yield (left) and % fat (right); ANN1: Artificial neural network prediction for the assorted data; A: Predicted output; T: Observed data

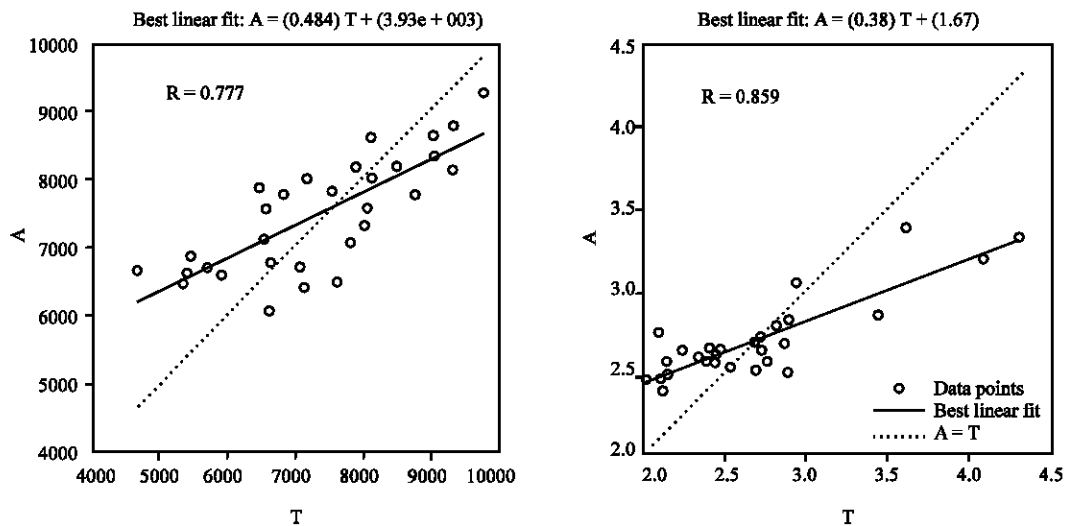


Fig. 4: Regressions of ANN2 predicted on the observed data for second parity milk yield (left) and % fat (right); ANN2: Artificial neural network prediction for the unsorted data; A: Predicted output; T: Observed data

As the training concept suggests, prediction of unseen records by a neural network improves when similar cases have been included during training (Salehi *et al.*, 1998). Thus, selection of a proper sample is very important for training an ANN. For example, if the training sample was selected randomly (in an adequate number) from the total population, the predictions can be extended to the total population. However, if the majority of the sample was from the upper or lower extremes of the population, the network will do better predictions for that part of data and the predictions for the other parts may not be accurate. Pre-processing data (e.g., standardization and normalization) may lead to an improvement in the

learning process of ANNs (Stein, 1993) which helps neural networks to predict better.

Relationships (regressions) between the observed and predicted records by ANNs are shown in Fig. 3 and 4. As a general rule, as the angle between the best linear fit and the 45° hypothetical line becomes smaller, ANN predictions become more reliable due to the closeness of the predicted to the observed data. The results of the current study support the efficiency of ANNs to predict second parity yields, based on available information on the first parity. It means that ANNs can be interesting alternatives to the traditional methods for predicting successive productions using previous

records. The results were in agreement with the findings of Lacroix *et al.* (1995) on the prediction of 305 days full parity records from part lactation records. Also, Yang *et al.* (1999) supported the findings that ANNs using assorted data, have an important advantage to the conventional methods of prediction.

### CONCLUSIONS

The major use of any predictive process is to support accurate decisions which are dependent on prior knowledge of the possible outcome and the outcome of this study showed that generally, neural networks have the ability to predict milk yield and fat percentage with high accuracy. The efficiency of ANNs will be more improved when samples and variables which are more relevant to the output variables are used. The results also showed that, first parity information can be used in prediction of milk yield and fat percentage of the second parity with high correlations with the observed data. Then, it could be concluded that ANN has a good potential to be used in prediction of subsequent records of dairy cows, in order to setup selection programs for increasing genetic potential of dairy herds and it is a good support system for dairymen for decision making. A flexibility of this method is that it can be further developed for health, fertility, lifetime and other economical traits in dairy industry. The authors also suggest an investigation on the reliability and efficiency of new constructed neural networks comparing to the standard classical BLUP (best linear unbiased prediction) method.

### ACKNOWLEDGMENT

The authors acknowledge Isfahan University of Technology for financial support of P. Hosseini's M.Sc. Thesis.

### REFERENCES

- Adamczyk, K., K. Molenda, J. Szarek and G. Skrzyński, 2005. Prediction of bulls slaughter from growth data using artificial neural network. *J. Cent. Eur. Agric.*, 6: 133-142.
- Brethour, J.R., 1994. Estimating marbling score in live cattle from ultrasound images using pattern recognition and neural network procedures. *J. Anim. Sci.*, 72: 1425-1432.
- Brons, A., G. Rabatel, F. Fose, F. Sevila and C. Touzet, 1993. Plant grading by vision using neural networks and statistic. *Comput. Electron. Agric.*, 9: 25-39.
- Dayhoff, J.E., 1990. *Neural Network Architectures; An Introduction*. Van Nostrand Reinhold Co., New York, USA.
- Elizondo, D.A., R.W. McClendon and G. Hoogenboom, 1994. Neural network models for predicting flowering and physiological maturity of soybean. *Trans. ASAE.*, 37: 981-988.
- Ganesan, V., N. Joshi and K. Muthukumarappan, 2004. Correlating calcium concentration and rheological characteristics of Mozzarella cheese using neural network. *ASAE Annual Meeting*, Paper No. 046044.
- Grzesiak, W., P. Błaszczyk and R. Lacroix, 2006. Methods of predicting milk yield in dairy cows-predictive capabilities of wood's lactation curve and Artificial Neural Networks (ANNs). *Comput. Electron. Agric.*, 54: 69-83.
- Hagan, M.T. and M. Menhaj, 1994. Training feed forward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks*, 5: 989-993.
- Kominakis, A.P., Z. Abas, I. Maltaris and E. Rogdakis, 2002. A preliminary study of the application of artificial neural networks to prediction of milk yield in dairy sheep. *Comput. Electron. Agric.*, 35: 35-48.
- Korthals, R.L., G.L. Hahan and J.A. Nienaber, 1994. Evaluation of neural networks as a tool for management. *Trans. ASAE*, 37: 1295-1299.
- Kwonhoh, H., Q. Zhang and S. Han, 2004. A neural network model of nitrogen stress assessment using multispectral corn nitrogen deficiency sensor. *ASAE Annual Meeting*, Paper No. 041132.
- Lacroix, R., K.M. Wade, R. Kok and J.F. Hayes, 1995. Prediction of cow performance with a connectionist model. *Trans. ASAE.*, 38: 1573-1579.
- Lacroix, R., F. Salehi, X.Z. Yang and K.M. Wade, 1997. Effects of data preprocessing on the performance of artificial neural networks for dairy yield prediction and cow culling classification. *Trans. ASAE.*, 40: 839-846.
- López-Benavide, M.G., S. Samarasinghe and J.G.H. Hickford, 2003. The use of artificial neural networks to diagnose mastitis in dairy cattle. *The International Joint Conference on Neural Networks (IJCNN)*: (CD-ROM).
- MATLAB, 2006. *MATLAB Guide for Personal Computers*. Version 7.3 Edn. MATLAB Institute, USA.
- Md Saad, M.H., M.J. Mohd Nor, F.R.A. Bustami and R. Ngadiran, 2007. Classification of heart abnormalities using artificial neural network. *J. Applied Sci.*, 7: 820-825.
- Molenda, K., Z. Gil and J. Żychlińska, 2001. Conception of computer system support in detection of pregnancy, oestrus and mastitis based on milk temperature. 52nd Annual Meeting of EAAP, Budapest, Hungary.

- Parmer, R.S., R.W. McClendon, G. Hoogenboom, P.D. Blankenship, R.J. Cole and J.W. Dorner, 1997. Estimation of aflatoxin contamination in pre-harvest peanuts using neural networks. *Trans. ASAE.*, 40: 809-813.
- Sablani, S.S., H.S. Ramaswamy and S.O. Prasher, 1995. A neural network approach for thermal processing applications. *J. Food Proc. Preserv.*, 19: 283-301.
- Salehi, F., R. Lacroix and K.M. Wade, 1998. Effects of learning parameters and data presentation on the performance of back-propagation networks for milk yield prediction. *Trans. ASAE.*, 41: 253-259.
- Salehi, F., R. Lacroix and K.M. Wade, 2000. Development of neuro-fuzzifiers for qualitative analyses of milk yield. *Comput. Electron. Agric.*, 28: 171-186.
- Stein, R., 1993. Preprocessing data for neural networks. *AI Expert*, 8: 32-37.
- Wildberger, A.M., 1990. Neural networks as a modeling tool. *AI and Simulation. Theory and Application*, San Diego, Calif., pp: 65-74.
- Yang, X.Z., R. Lacroix and K.M. Wade, 1999. Neural detection of mastitis from dairy herd improvement records. *Trans. ASAE.*, 42: 1063-1072.