



Mastitis detection in Murrah buffaloes with intelligent models based upon electro-chemical and quality parameters of milk

Indu Panchal*, I.K. Sawhney, A.K Sharma¹, M.K. Garg² and A.K. Dang³

ICAR-National Dairy Research Institute,
Karnal-132 001, Haryana, India.

Received: 10-12-2015

Accepted: 29-03-2016

DOI:10.18805/ijar.10773

ABSTRACT

In this paper, several connectionist models have been described to detect mastitis in Murrah buffaloes using milk parameters, viz., pH, electrical conductivity, temperature (udder, milk and skin), milk somatic cells, milk yield and dielectric constant. A total of 600 milk samples were collected from 100 lactating Murrah buffaloes; which were analysed for Somatic Cell Counts in milk. Accordingly, animals were classified into three categories, i.e., healthy, subclinical mastitis and clinical mastitis animals. These basal values were utilised for developing connectionist models to identify healthy versus mastitis animals. Also, Multiple Linear Regression (MLR) models were developed for comparing classification accuracy of proposed connectionist models using Root Mean Square Error (RMSE) technique. The connectionist models were found to be superior (RMSE = 0.01) as compared to MLR models (RMSE = 4.08). Hence, it is deduced that connectionist approach could be used as a suitable technique for detecting mastitis in Murrah buffaloes.

Key words: Connectionist models, Dairy, Error back propagation, Mastitis, Murrah buffaloes.

INTRODUCTION

Mastitis refers to inflammation of the mammary gland. The interior environment of the mammary gland is most favourable for the growth and multiplication of bacteria, which enter through the streak canal. Once entered inside the teat, if they are not eliminated, they start multiplying in the mammary gland and liberate toxins and induce leukocytes and epithelial cells to release chemo attractants. The dead and sloughed off mammary epithelial cells, in addition to the dead leukocytes, are secreted into the milk, resulting in high milk somatic cell counts (SCC) (Srivastava *et al.*, 2015).

Mastitis is the most costly disease in dairy animals and remains one of the main problems for the dairy-industry. This disease causes major economic losses to food industry as it influences the total milk output and modifies milk composition and technological usability (Schroeder, 2012). Routinely, the mastitis is detected by examining the milk somatic cell count (De *et al.*, 2010). However, assessment of SCC by using traditional methods is a time consuming process. On the other hand, costly equipments are required if the assessment is to be made using advanced methods. Thus, need of the hour is to find out an instant/cost-effective method for mastitis detection, which could be employed routinely on small scale dairy farms. The connectionist models are, therefore, proposed to predict mastitis in Murrah buffaloes.

The main advantage of the connectionist models over traditional methods is that they do not require clear, obvious mathematical information concerning the complex state of the problem being investigated. The connectionist model is defined as a complex system designed to emulate the human brain and nervous system; and is formed by artificial neurons or Processing Elements (PE's) connected together by means of various connection geometries. The features that distinguish connectionist models from other methods include its capacities for learning, generalisation and parallel processing. These features facilitate such advantages as speed, fault tolerance and efficiency (Mammadova and Keskin, 2015). The connectionist networks comprising multiple layers with nonlinear transfer function are capable to learn linear and nonlinear relationships between model inputs and outputs.

Connectionist models are being used increasingly in many fields of dairying (Atil and Akilli, 2015). Some studies have been conducted recently by various researchers that demonstrated application potential of connectionist models in dairying especially in predicting mastitis and ketosis risk in dairy animals. Mammadova and Keskin (2015) suggested connectionist models using input variables such as lactation rank (current lactation number), milk yield, electrical conductivity (EC), average milking duration and season to generate mastitis alerts. The cattle were judged

*Corresponding author's e-mail: indupanchal33@gmail.com. ¹Computer Centre and Division of Dairy Economics, Statistics & Management, ICAR – National Dairy Research Institute (Deemed University), Karnal-132001, Haryana, India. ²Department of Processing and Food Engineering, College of Agricultural Engineering, CCS Haryana Agricultural University, Hisar-125004, Haryana, India. ³Division of Dairy Cattle Physiology, ICAR – National Dairy Research Institute (Deemed University), Karnal-132001, Haryana, India.

healthy or infected based on SSC. Ankinakatte *et al.* (2013) developed connectionist models based upon EC, somatic cell scores, lactate dehydrogenase and milk yield for early detection of oncoming mastitis in dairy cattle with automated recorded data. Sun *et al.* (2010) described connectionist models to detect mastitis by automatic milking systems, using two mastitis indicators, *i.e.*, higher EC and lower quarter yield. Cavero *et al.* (2008) developed connectionist models to generate mastitis alerts using EC, milk production rate, milk flow rate and days in milk as input data. Hassan *et al.* (2007) developed unsupervised connectionist model based on changes in milk parameters associated with mastitis for detecting major and minor bacterial pathogens present in milk. Heald *et al.* (2000) developed a connectionist model to broadly classify bacterial causes of mastitis in dairy herds. Information describing herd management practices, quarter milk samples; and monthly data with moderate to high SSC was used to develop the connectionist model that discriminated among four categories of bacterial organisms (contagious, environmental, no significant growth, and other) associated with clinical and subclinical mastitis. Performance of the connectionist model was reported to be best *vis-à-vis* linear discriminant analysis, in herds with higher frequency of minor and contagious pathogens.

Evidently, review of relevant literature indicated that no study has been carried out to identify healthy *vs.* mastitis animals in indigenous Murrah buffaloes using connectionist models. Also, connectionist models developed in exotic cattle were based on individual parameter. Hence, in this paper, connectionist paradigm has been investigated on the basis of electro-chemical properties along with some other milk quality parameters of normal and mastitis milk for predicting/diagnosing mastitis in the Murrah buffaloes. These models would not only help in rapid detection of raw milk quality but also facilitate isolation of animals with subclinical mastitis for timely treatment. The classifying ability of connectionist models has been assessed *vis-à-vis* conventional regression models, which have also been developed in this study.

MATERIALS AND METHODS

The investigations were carried out on 100 lactating Murrah buffaloes set aside at Livestock Research Centre, Indian Council of Agriculture Research – National Dairy Research Institute, Karnal, India for a period of one year, *i.e.*, March, 2013 to February, 2014. A total of 600 milk samples were collected from these animals; and the same were screened for mastitis. The precision pH meter was used to measure pH while EC was determined with EC meter. The temperature of skin and udder of animals were recorded with infrared thermometer. Also, the temperature of milk was measured by normal thermometer. Dielectric constant was determined by dielectric constant meter. Further, SCC were determined by heating milk to 40°C in a water-bath,

which was kept at that temperature for 15 minutes before cooling it to 20°C by stirring gently. About 0.01 ml of milk was spread on a 1cm² (0.5 cm x 2 cm) area of a degreased microscopic slide and was dried in a horizontal position. SCC of milk samples were ascertained microscopically. Variables contained in the experimental dataset are given in Table 1.

Table 1: Variables used for models development.

Serial Number	Traits
1	Milk pH
2	Electrical Conductivity (mS/cm)
3	Udder Temperature (°C)
4	Milk Temperature (°C)
5	Skin Temperature (°C)
6	Milk Yield (kg)
7	Dielectric Constant

The values generated above were used for development of connectionist and Multiple Linear Regression (MLR) models to identify animals as healthy or mastitis Murrah buffaloes.

Connectionist models: Connectionist models reflects a system that is based on operations of biological neural networks; and, hence, can be defined as an emulation of biological neural systems. Connectionist models are at the forefront of computational systems designed to produce, or at least mimic intelligent behaviour. Connectionist models fall under Cognitive Science and have been founded on diverse sources. It can be used for inquisitiveness of mankind to understand and emulate human brain; copying human abilities like speech and the use of language; pattern recognition, modelling and prediction (Schroeder, 2012).

A connectionist model comprises three main components, *viz.*, i) Topology – organisation of connectionist network consisting of several layers and how these layers are connected to each other; ii) Learning – the algorithm that facilitates acquisition of knowledge from data within the network; and iii) Recall – retrieving the stored knowledge from the network. The primary configuration of a connectionist model consists of PE's that are grouped into layers. Connectionist network consists of an input layer, one or more hidden layer(s) and an output layer. Refere Mammadova and Keskin (2015) for further pedagogic details about connectionist models (also known as artificial neural networks).

Multiple linear regression models: Multiple linear regression models are the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data. Every value of the independent variable x is associated with a value of the dependent variable y . The population regression line for p explanatory variables $x_1, x_2, \dots, x_p$ is defined to be m .

The line, $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$ describes how the mean response m_y changes with the explanatory variables. The observed values for y vary about their means m_y and are assumed to have the same standard deviation s . The fitted values $\beta_0, \beta_1, \dots, \beta_p$ estimate the parameters $b_0, b_1, \dots, b_p$ of the population regression line.

Since the observed values for y vary about their means μ_y , the multiple regression models include a term for this variation. In simple words, the model is expressed as DATA = FIT + RESIDUAL, where the term "FIT" represents the expression $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$; the "RESIDUAL" term represents the deviations of observed values y from their means μ_y , which are normally distributed with mean 0 and variance σ . The notation for the model deviations is $\hat{a}$. Formally, the model for multiple linear regressions, given n observations, is

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \text{ for } i = 1, 2, \dots, n.$$

Models development: For development of the models, the raw data were divided into two subsets, viz., training set and test set. The training set was used for model training while the test set was used for model validation. Three data partitioning schemes such as 70:30, 80:20 and 90:10 were explored for model development. Each training experiment was conducted ten times with different combinations of the parameters such as number of neurons and transfer function in the hidden layer, data portioning scheme, error goal, learning rate, epochs, etc. For the model selection, the architecture of connectionist model was decided by 'trial and error' technique. The MATLAB programming environment was used for training and simulation experiments. SAS 9.3 was used for developing MLR models.

Model performance analysis: The Root Mean Square Error (RMSE) technique was used to examine performance of connectionist models and MLR models. The prediction error and RMSE have been calculated using the following formulae:

$$\text{Prediction Error \%} = \frac{\text{Actual value} - \text{Predicted value}}{\text{Actual value}} \times 100$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{Prediction Error \%})^2}$$

RESULTS AND DISCUSSION

Connectionist model performance: Performance results of feed-forward sigmoid connectionist models based on Error Back Propagation (EBP) learning algorithm with Bayesian regularisation mechanism developed in this paper with different data partitioning schemes (*i.e.*, 70:30, 80:20 and 90:10) to detect mastitis milk in Murrah buffaloes, are presented in Tables 2 through 4.

The incoming data were processed by nonlinear function at hidden layer and by linear function at the output layer so as to get the net output. A comprehension search of connectionist model architecture was performed by varying the number of neurons between 2 and 6 in the hidden layer. Either log-sigmoid function or tangent-sigmoid function was employed on hidden layer neurons as transfer function. The connectionist network was best trained with either 500 or 1000 epochs; learning rate as 0.01 and error goal as 0.001, which were determined empirically. The network error was computed by comparing the network output with the target output, and optimal network was selected based on minimum RMSE. A code written in MATLAB language was used for

Table 2: Feed-forward sigmoid connectionist models based on EBP learning algorithm with Bayesian regularisation mechanism and data partitioning scheme as 70:30 to detect healthy vs. mastitis Murrah buffaloes.

Transfer function on hidden layer neurons	Epochs	RMSE (%)				
		Number of neurons in hidden layer				
		2	3	4	5	6
Log-sigmoid	500	0.172	0.191	0.148	0.106	0.176
	1000	0.168	0.151	0.227	0.185	0.162
Tangent-sigmoid	500	0.168	0.196	0.149	0.183	0.224
	1000	0.168	0.186	0.149	0.166	0.234

Table 3: Feed-forward sigmoid connectionist models based on EBP learning algorithm with Bayesian regularisation mechanism and data partitioning scheme as 80:20 to detect healthy vs. mastitis Murrah buffaloes.

Transfer function on hidden layer neurons	Epochs	RMSE (%)				
		Number of neurons in hidden layer				
		2	3	4	5	6
Log-sigmoid	500	0.150	0.170	0.130	0.101	0.144
	1000	0.150	0.152	0.165	0.139	0.141
Tangent-sigmoid	500	0.154	0.162	0.150	0.137	0.141
	1000	0.154	0.159	0.166	0.155	0.060

Table 4: Feed-forward sigmoid connectionist models based on EBP learning algorithm with Bayesian regularisation mechanism and data partitioning scheme as 90:10 to detect healthy vs. mastitis Murrah buffaloes.

Transfer function on hidden layer neurons	Epochs	RMSE (%)				
		Number of neurons in hidden layer				
		2	3	4	5	6
Log-sigmoid	500	0.124	0.014	0.015	0.010	0.014
	1000	0.015	0.016	0.024	0.073	0.061
Tangent-sigmoid	500	0.136	0.125	0.067	0.107	0.078
	1000	0.148	0.176	0.144	0.028	0.040

carrying out the training and simulation experiments. The least value of RMSE was found with data partitioning scheme as 90:10 with 5 neurons and log-sigmoid transfer function in hidden layer and 500 epochs. Schematic diagram of connectionist model to identify healthy vs. mastitis Murrah buffaloes has been depicted in Fig. 1.

Multiple linear regression model performance: The Multiple linear regression models were also developed with the same data partitioning schemes as used for constructing the connectionist models described in previous section, *i.e.*, 70:30, 80:20 and 90:10. The MLR equations thus developed for Murrah buffaloes are shown in Table 5.

The accuracy of fit was checked by calculating the RMSE between experimental and predicted values of SCC. The lower the value of RMSE the better was the goodness of fit. Generally, a good description of data is considered, on an average, with RMSE to be smaller than 7%. The mastitis detection accuracy of MLR models varied between 4.08% and 22.90%.

Connectionist models *vis-à-vis* conventional regression Models: The comparative performance of connectionist models *vis-à-vis* MLR models developed in this paper was made in terms of RMSE on test set (Table 6). In Murrah buffaloes, the least value of RMSE (0.010) was found with

Table 5. Multiple linear regression equations to identify healthy vs. mastitis Murrah buffaloes.

Data Partitioning Scheme	MLR Equation
70:30	$SCC = 0.059pH + 5.53EC + 0.034UT + 0.179MT + 0.063ST + 0.29MY + 0.65DC - 73.46$
80:20	$SCC = 0.126pH + 5.50EC - 0.04UT + 0.21MT - 0.025ST + 0.25MY + 0.53DC - 61.15$
90:10	$SCC = 0.198pH + 5.41EC + 0.079UT + 0.263MT - 0.035ST + 0.20MY + 0.4433DC - 60.25$

SCC: Somatic Cell Counts; EC: Electrical Conductivity; UT: Udder Temperature; MT: Milk Temperature; ST: Skin Temperature; MY: Milk Yield; DC: Dielectric Constant.

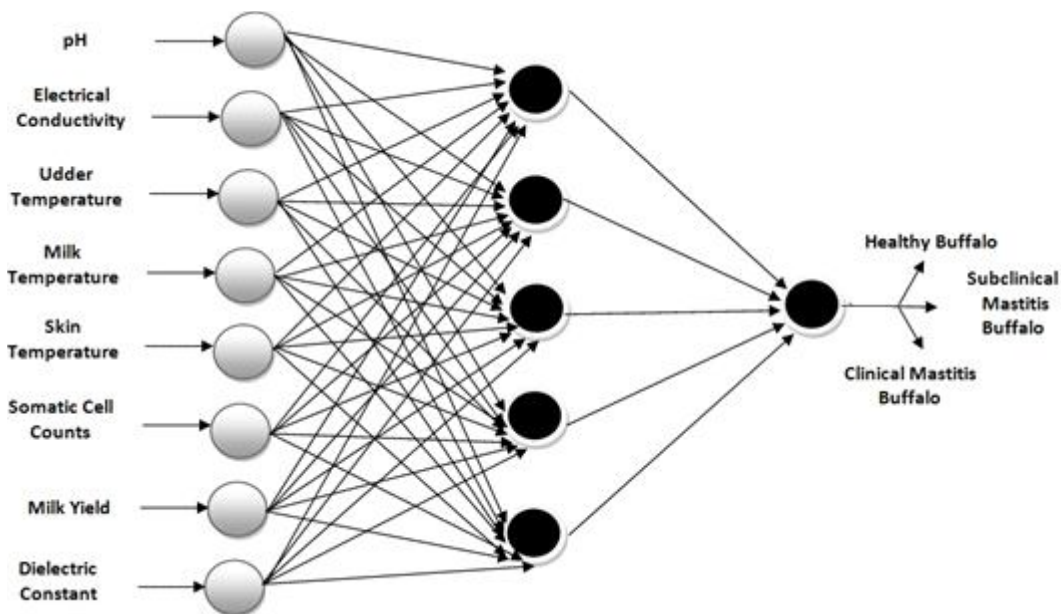
**Fig 1:** Schematic diagram of connectionist model to identify healthy vs. mastitis Murrah buffaloes.

Table 6: Comparative performance of connectionist models vis-à-vis MLR models with different data partitioning schemes, to identify healthy vs. mastitis Murrah buffaloes.

Data Partitioning Scheme	RMSE (%)	
	Connectionist model	MLR model
70:30	0.106	22.90
80:20	0.060	4.72
90:10	0.010	4.08

data partitioning scheme as 90:10 whereas that of MLR models was found to be 4.08 for the same data partitioning. Hence, it is concluded that the proposed connectionist classifier is efficient and suitable technique for identifying healthy and mastitis Murrah buffaloes.

CONCLUSION

In this paper, several connectionist models have been developed and validated to identify healthy vs. mastitis Murrah buffaloes on the basis of electro-chemical parameters along with milk quality parameters of normal and mastitis milk. The error back propagation training algorithm with Bayesian regularisation scheme and several combinations of different values of network parameters was empirically investigated. The performance of connectionist models was compared to that of classical multiple linear regression models. The comparative analysis of the results thus obtained revealed that connectionist paradigm was more superior and efficient than classical regression method to diagnose healthy vs. mastitis Murrah buffaloes.

REFERENCES

- Ankinakatte, S.A., Norberg, E., Lovendahl, P., Edwards, D. and Hojsgaard, S. (2013). Predicting mastitis in dairy cows using neural networks and generalized additive models: a comparison. *Comput. Electron. Agr.* **99**: 1-6.
- Atil, H. and Akilli, A. (2015). Investigation of dairy cattle traits by using artificial neural networks and cluster analysis. In: *Proceedings of the 7th International Conference on Information and Communication Technologies in Agriculture, Food and Environment (HAICTA 2015)*, Kavala, Greece, September 17-20, pp. 681-690.
- Cavero, D., Tolle, K.H., Buxade, C. and Krieter, J. (2008). Mastitis detection in dairy cows by application of neural networks. *Livest. Sci.* **114**: 280-286.
- De, K., Mukherjee, J., Prasad, S. and Dang, A.K. (2010). Effect of different physiological stages and managerial practices on milk somatic cell counts of Murrah buffaloes. In: *Proceedings of the 9th World Buffalo Congress*, Buenos Aires, Argentina, April 25-28, pp. 549-551.
- Hassan, K.J., Samarasinghe, S. and Lopez-Benavides, M.G. (2007). The use of neural networks to detect minor and major pathogens that cause bovine mastitis. In: *Proceedings of the New Zealand Society of Animal Production*, **67**: 215-219.
- Heald, C.W., Kim, T., Sischo, W.M., Cooper, J.B. and Wolfgang, D.R. (2000). A computerized mastitis decision aid using farmbased records: An artificial neural network approach. *J. Dairy Sci.* **83**:711-720.
- Lopez-Benavides, M.G., Samarasinghe, S. and Hickford, J.G.H. (2003). The use of artificial neural networks to diagnose mastitis in dairy cattle. In: *Proceedings of the International Joint Conference on Neural Networks*, July 20-24, **1**: 582 - 585.
- Mammadova, N.M. and Keskin, I. (2015). Application of neural network and adaptive neuro-fuzzy inference system to predict subclinical mastitis in dairy cattle. *Indian J. Anim. Res.* **49**: 671-679. www.arccjournals.com/uploads/articles/19671679B273.pdf.
- Schroeder, J.W. (2012). Mastitis control programs: Bovine mastitis and milking management. NDSU Extension Service publication: AS1129, North Dakota State University, Fargo, North Dakota. www.ag.ndsu.edu/pubs/ansci/dairy/as1129.pdf.
- Srivastava, A.K., Kumaresan, A., Manimaran, A. and Prasad, S. (2015). Mastitis in dairy animals - An update. Satish Serial Publishing House, New Delhi, India. ISBN-10: 9384053066.
- Sun, Z., Samarasinghe, S. and Jago, J. (2010). Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks. *J. Dairy Res.* **77**:168-75.