Wiadomości Zootechniczne, LV (2017), 3: 56–63

Estimating the effect of some milk components on urea concentration using artificial neural networks

Edyta Bauer[1], Justyna Żychlińska-Buczek[2]

University of Agriculture in Krakow, al. Mickiewicza 24/28, 32-120 Kraków, Poland
[1]Department of Genetics and Animal Breeding, [2]Department of Cattle Breeding

The aim of the study was to use artificial neural networks (ANN) to assess the impact of subsequent lactations on the urea content in milk. Milk components were used in the observation, such as: the percentage content of fat, protein, the fat-to-protein and fat-to-lactose ratio. The material for the study included data obtained from trial milkings in a production farm in the Lower Silesia Voivodeship. There were Polish Holstein-Friesian (HF) cows of black- and red-white varieties located on the farm. The milk yield on the farm was ± 1200 L for lactation. The performed analyzes showed that the influence of selected milk components in particular lactations changes on the basis of the selected ANN model. The obtained results may confirm the effect of subsequent lactations on the content of urea and selected components in milk. The ANNs used in the study can be used as a tool to support modeling and identification of non-linear and difficult to determine interdependencies between factors influencing their course.

Artificial neural networks (ANNs) are a tool that is increasingly used in learning about complex biological processes. The basis for developing the network includes training algorithms (Korbicz et al., 1994; Kosiński, 2002; Rzewulska & Strabel, 2013 a,b), enabling the design of an appropriate network structure, and the selection of parameters for this structure, suited to the problem being solved. ANN simulate the central nervous system function in living organisms in a simplified way. Modeling of information processing processes allows to solve difficult or even impossible problems with standard methods. Biological processes are characterized by high complexity and non-linearity as well as interdependencies between factors affecting their course which are difficult to determine. ANNs are used to assist in supporting, modeling and learning about the processes taking place in a broad sense of agricultural engineering (Boniecki, 2004, 2005; Kosiński, 2014). The potential SNN application is constantly expanding. This is possible due to their basic property, which is the use of the training process. Network's training consists of automatic (according to the appropriate algorithm) selection of weight values at which the network will solve the given task as well as possible. The training process is carried out on the basis of data in the network. The advantage of ANN is the distributed nature of information processing with a large number of parameters (network and training algorithm) with no strict rules to estimate their value. These properties result in resistance to damage (elimination of a significant part of the neurons) or less clear data which occurs in case of empirical data (Rzewulska & Strabel, 2013 b).

Properly balanced in terms of protein and energy content, the feed dose for dairy cows is a basic condition for optimal lactation and proper milk composition (Jamroz & Potkański, 2004). The high-yielding cows are an especially sensitive feeding group, in case of which attention should be paid to the feeding dose share of feeds providing two basic components, i.e. protein and energy, and their interdependencies. The deficiencies of these components (Johnson &

Young, 2003) can be found particularly in cows at the peak lactation period, when the animal's nutritional needs exceed the possibility of taking appropriate components from the feed, and the resulting deficit must be covered from the reserves of the organism. The urea content in milk depends e.g. on the intake of general protein and energy, and on the relative ratio of these components in the feeding dose (Hojman et al., 2004). The amount of protein as well as urea in milk is an indicator of a properly balanced food dose (Hojman et al., 2004). Many factors may affect the urea content, including: season, month, lactation phase, milking day, as well as time of day and lactation number (Arunvipas et al., 2003; Guliński et al., 2008; Fatehi et al., 2012). According to various authors, the heritability of urea content in milk results from individual variability (Rajala-Schultz & Saville, 2003; Bastin et al., 2009; Satoła & Ptak, 2016).

The aim of this study was to design an appropriate network structure (ANS) and to select the parameters of this structure to investigate the relation between the urea content in milk using ingredients such as: % protein content, % fat content, ratio of fat to protein and fat to lactose in milk of Polish Holstein-Friesian (HF) cows of the black- and red-white varieties and subsequent lactations using artificial neural networks with radial base functions (RBF).

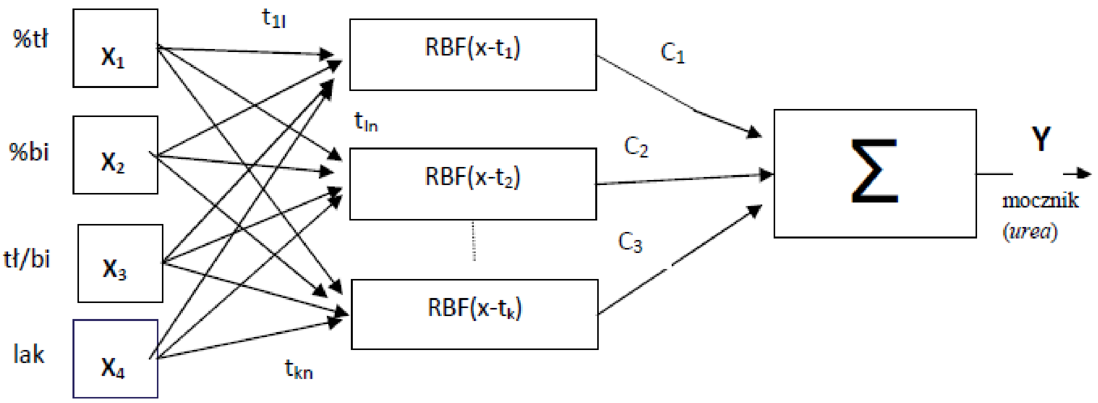
Materials and methods

In order to conduct studies on the impact of successive lactations on the urea content in milk, the Radial Basis Functions simulator from the resources of the *Statistica12®* package was used. The selection of ANN for the analysis of data was determined by their feature, the analysis network sensitivity to individual input variables. The input variables were % protein, % fat, fat to protein ratio and % lactose. Sensitivity analysis was carried out to obtain information on the relations contained in the given data set (Rajala-Schultz & Saville, 2003). The usefulness of the sensitivity analysis to assess the impact of the lactation number on the urea content in milk was used (Kosiński, 2002; Niżewska et al., 2007; Guliński et al., 2008; Fatehi et al., 2012).

In order to generate a set of appropriate RBF topologies of neural network, an effective procedure implemented in *Statistica v.12®* software in the form of an automatic neural network designer was used. This tool allows to reduce the search time of the proper neural networks, training on the basis of given set of variables.

Sensitivity analysis was carried out based on the composition of milk in subsequent lactations. For this purpose, the percentages of protein, fat, fat, fat, fat-to-protein ratio, and lactose were used. The data included 2590 trial milkings, obtained from 370 Polish Holstein-Friesian cows, black- and red-white varieties. The cows were divided depending on the lactation number (I – 108 cows, II – 124 cows, III – 138 cows). There were 265 cows of the black and white variety and 105 cows of the red and white variety. The analysis involved the use of milk ingredients from a minimum of 7 trial milkings per lactation from one cow. The animals were kept in a free-standing system in the production farm. The herd was divided into two production groups. The main group were cows with a milk yield of ± 20 kg of milk per day. The animals were fed with full dose rations (Total Mix Ration). Trial milkings were performed in the first, second and third lactations in the period from April to December 2013 by a specialist from the Polish Federation of Cattle Breeders and Milk Producers in accordance with the standards.

The urea content in milk was used for interpretation (as the output variable). The main analyzed input variables included milk components, such as: percentage content of fat, protein, lactose and fat to protein ratio. The trial milkings included in the analysis were quantitatively different in composition due to the random selection of study groups. Accepted course of action in the analysis is shown in Figure 1. The selection of fixed parameters for the basic functions is performed randomly with a uniform distribution. The selected Gaussian radial function assumes the value of the standard deviation of functions dependent on the distribution of randomly selected centers. The distance between the centers is the same for each basic function.



t_n – zestaw wag j-tego neuronu warstwy ukrytej – *weights for the j_{th} hidden layer neuron;*
x_1 -x_4 – zmienne wejściowe (% tłuszczu, % białka, tłuszcz/białko, laktoza) – *input variables (%fat, %protein, fat/protein, lactose);*
x_i -t_i – neuron radialny (RBF) – *radial neuron (RBF);*
c_n – waga pomiędzy j-tym neuronem warstwy radialnej a neuronem warstwy wyjściowej – *weight between the j_{th} radial and output layer neurons;*
y – wartość wyjściowa neuronu wyjściowego (mocznik) – *input value of the input neuron (urea)*

Fig. 1. Neural network topology RBF model (Radial Basis Function)

Results and discussion

The random nature of the network training process means that neural networks developed for modeling the same process may differ in terms of error values (Rajala-Schultz & Saville, 2003). Therefore, when selecting a network, it is necessary to adopt the best model. In this study, it was assumed that the criterion for choosing the best model of the neural network would be the elements of the training set divided into groups of similar elements. In the RBF (Radial Basis Function) network, the division of radial neuron weights by k-means used to create k-clusters by specifying modal elements for each of them, representing the agreed middle of the entire group in the space of properties (Rajala-Schultz & Saville, 2003). In order to find the best locations for the central points of each cluster, random localization is assumed at the beginning, and then they are perfected to optimally match each pattern to the input data cluster whose center is the closest to the pattern (Rajala-Schultz & Saville, 2003; Boniecki, 2005).

When using the linear model, the error function was based on the sum of squares (Root Mean Squared). RMS error is a total error made by the network on a certain set of data. The error equals the sum of the squares of differences between the values predicted (by model) and real (observed) ones obtained in case of the perceptron three-layer BRF network, taught by the universal approximation of any continuous function with sufficient accuracy. The output layer included simple linear (weighted) summation.

Network training occurs as a problem of approximation, the best fit (reconstruction) of the hypersurfaces to the training data.

Training of output neurons is done through a single output and determination of the weight vector. The choice of ANN for the analysis of data was determined by their characteristic, in the form of analyzing the network's sensitivity to particular input variables. The sensitivity analysis used in networks allows allows to distinguish important variables from those that do not contribute much as a result of the network's operation. Indicates variables that can be omitted without loss of network quality, and key variables which must not be omitted. Input variables are usually not dependent.

Table 1. Parameters generated by RBF networks for I, II and III lactation

Network tests for dataset **Lactation I**							
	TYPE	quality (u)	quality (t)	quality (w)	algorithm	activations (hidden layer)	activations (input layer)
1	RBF 4-11-1	0.30	0.14	4.69	RBF	Gauss	linear
2	RBF 4-11-1	0.37	0.22	1.03	RBF	Gauss	linear
3	RBF 4-11-1	0.37	0.41	5.06	RBF	Gauss	linear
Network tests for dataset **Lactation II**							
	TYPE	quality (u)	quality (t)	quality (w)	algorithm	activations (hidden)	activations (input layer)
1	RBF 4-11-1	0.33	0.26	0.13	RBF	Gauss	linear
2	RBF 4-11-1	0.32	0.21	0.21	RBF	Gauss	linear
3	RBF 4-11-1	0.34	0.06	0.06	RBF	Gauss	linear
Network tests for dataset **Lactation III**							
	TYPE	quality (u)	quality (t)	quality (w)	algorithm	activations (hidden)	activations (input layer)
1	RBF 4-11-1	0.72	0.63	0.47	RBF	Gauss	linear
2	RBF 4-11-1	0.72	0.48	0.43	RBF	Gauss	linear
3	RBF 4-11-1	0.72	0.59	0.37	RBF	Gauss	linear

RBF – Radial Basis Function; u – training test; t – test set; w – validation test.

The generated neural networks had a structure of 4:11:1 type – meaning there was 1 hidden layer, containing 11 neurons having an exponential function as an activation function. The number of input neurons was 4 (% fat, % protein, fat to protein ratio, % lactose), with 1 output neuron (urea). The training set, test set, and cross-validation are used to evaluate the quality of network qualifications. The training data is then divided into two separate subsets: the training sample (u) and the test sample (t). The network is trained by means of a training sample and the quality of the network is examined by means of a test sample. The structure of generated and selected networks is presented in Tab. 1. The three columns in Tab. 1 define the network quality for the training, test and validation sets. This quality is expressed by means of a linear correlation coefficient between the values of dependent variables and the network output.

Based on selected models of the RBF network, sensitivity analysis of input variables was carried out. The sensitivity analysis is the measure of the neuronal model for input variables, appearing as the ratio of errors, which indicates an increase in the total error made by the model when the analyzed output variable is not taken into account. The greater the ratio of errors, the greater the effect of the variable on the final result of the network.

If the error ratio for a given variable is equal to or lower than 1, then the variable has no impact on the quality of training and the result of the network (Rzewulska & Strabel, 2013 a). Tables 2, 3, and 4 provide sensitivity analysis for the obtained networks and individual lactations.

Table 2. Sensitivity analysis for lactation I

Sensitivity analysis for training, test and validation sets, lactation I				
networks	% lactose	% protein	fat/protein	% fat
RBF 4-11-1	1.00	0.99	0.99	0.99
RBF 4-11-1	0.99	1.00	0.99	0.99
RBF 4-11-1	1.02	1.02	1.02	1.01

Table 3. Sensitivity analysis for lactation II

Sensitivity analysis for training, test and validation sets, lactation II				
networks	% lactose	fat/protein	% protein	% fat
RBF 4-11-1	1.11	1.08	1.03	1.04
RBF 4-11-1	1.06	1.01	1.06	1.02
RBF 4-11-1	1.08	1.15	1.06	1.07

Table 4. Sensitivity analysis for lactation III

Sensitivity analysis for training, test and validation sets, lactation III				
networks	% lactose	% fat	% protein	fat/protein
RBF 4-11-1	1.25	1.16	1.18	1.07
RBF 4-11-1	1.15	1.12	1.07	1.07
RBF 4-11-1	1.27	1.17	1.16	1.07

RBF – Radial Basis Function; % fat – percentage of fat; % protein – percentage of protein; % fat/protein – fat to protein percentage; % lactose – percentage of lactose.

Based on the results of the sensitivity analysis included in Tables 2, 3 and 4, it was found that all milk components used from the test milkings affected the urea content in the milk. The results of the sensitivity analysis for lactation I indicate the possibility of continuing the modeling of the network with the omission of input variables, such as: percentage fat content and fat to protein ratio. Lactation I differs in terms of the results of the sensitivity analysis from lactation II and III. The literature contains publications confirming the differences between lactations (Jonker et al., 1999; Fatehi et al., 2012; Rzewulska & Strabel, 2013 b). The high urea content in milk indicates the ineffective use of total protein which increases the costs related to cow feeding and pollution (Rzewulska & Strabel, 2013 a).

The relations between the content of urea and subsequent lactation was the subject of many studies (Guliński et al., 2008; Fatehi et al., 2012).

In their studies, the authors found that the urea content in milk from the first two lactations was higher than from subsequent lactations. The results of studies conducted by many other authors indicate urea level fluctuations with diversified milk yield (Johnson & Young, 2003; Guliński et al., 2008; Sawa et al., 2010; Fatehi et al., 2012; Rzewulska & Strabel, 2013 a,b). The results of the sensitivity analysis indicate a significant effect of fat in the third lactation. As the urea content increases, the fat content of high-yielding herds increases. This relationship was confirmed by studies (Niżewska et al., 2007). Researches conducted by other authors indicated that as the content of fat increased, the content of urea in milk decreased (Guliński et al., 2008).

The results described above also show the effect of protein on the urea content in the first lactation, while our own studies prove that the results obtained in subsequent lactations affect lactose and fat/protein ratios in sensitivity analysis. The dependence between protein content increase in milk and the urea content has been confirmed by studies of many authors (Jonker et al., 1999; Godden et al., 2001; Fatehi et al., 2012; Rzewulska & Strabel, 2013 a), who focus on the increase in protein and urea content within particular lactations. The literature confirms that urea content increases with higher milk yield (Godden et al., 2001; Arunvipas et al., 2003).

In conclusion, one may say that the urea content in cow milk is affected by the lactose, protein, fat and fat-to-protein ratio. The content of these components varies depending on the subsequent lactation. The effect of lactose and protein on the urea content was significant in the sensitivity analysis for first lactation. In case of the second and third lactation, the influence of the lactose percentage was found. The sensitivity analysis of the generated network models demonstrated the effects of the lactation number on the urea content in case of selected milk components used as input neurons in the network.

Conclusions

The initial state of the network affects its training capabilities in a certain way. The use of an increased number of input neurons (input variables – milk components) or hidden ones (radial neurons) can facilitate or hinder the learning process of the network. This is done by increasing the diversity in the final weights. Required degree of network construction complexity (number of hidden neurons, number of weight reduction) is related to the difficulty of the problem. ANNs seem to be an appropriate tool for further detailed observations regarding the urea content in the milk of high yielding cows, using a wider base of dependent variables. It seems advisable to use modern methods of construction and use of neural models in the process of analyzing the impact of non-nutritional factors on the chemical composition of the obtained milk.

References

Arunvipas P., Dohoo I.R., Vanleewen J.A., Keefe G.P. (2003). The effect of non-nutritional factors on milk urea nitrogen levels in dairy cows in Prince Edward Island, Canada. Prev. Vet. Med., 59: 83–93.

Bastin C., Laloux L., Gillon A., Miglior F., Soyeurt T.H., Hammami H., Bertozzi C., Gengler N. (2009). Modeling milk urea of Walloon dairy cows in management perspectives. J. Anim Sci., 92: 3529–3540.

Boniecki P. (2004). Sieci neuronowe typu MLP oraz RBF jako komplementarne modele aproksymacyjne w procesie predykcji plonu pszenżyta. J. Res. Appl. Agr. Eng., 49 (1): 28–33.

Boniecki P. (2005). Wykorzystanie technik neuronowych w praktyce rolniczej. J. Res. Appl. Agr. Eng., 50 (2): 10–14.

Fatehi F., Zali A., Honarvar M., Dehghan-Banadaky M., Young A.J., Ghiasvand M., Eftekhari M. (2012). Review of the relationship between milk urea nitrogen and days in milk, parity, and monthly temperature mean in Iranian Holstein cows. J. Dairy Sci., 95: 5156–5163.

Godden S.M., Lissemore K.D., Kelton D.F., Leslie K.E., Walton J.S., Lumsden J.H. (2001). Factors associated with milk urea concentrations in Ontario dairy cows. J. Dairy Sci., 84: 107–114.

Guliński P., Młynek K., Salamończyk E. (2008). Zmiany zawartości mocznika w mleku w zależności od wybranych czynników środowiskowych. Med. Weter., 64: 465–468.

Hojman D., Kroll O., Adin G., Gips M., Manochi B., Ezra E. (2004). Relationships between milk urea and production, nutrition and fertility traits in Israeli dairy herds. J. Dairy Sci., 87: 1001–1011.

Jamroz D., Potkański A. (2004). Żywienie zwierząt i paszoznawstwo. Podstawy szczegółowego żywienia zwierząt. Wyd. Nauk. PWN, Warszawa.

Johnson R.G., Young A.J. (2003). The association between milk urea nitrogen and DHI production variables in Western commercial dairy herds. J. Dairy Sci., 81: 2681–2692.

Jonker J.S., Kohn R.A., Erdman R.A. (1998). Using milk urea nitrogen to predict nitrogen excretion and utilization efficiency in lactating dairy cows. J. Dairy Sci., 81: 2681–2692.

Jonker J.S., Kohn R.A., Erdman R.A. (1999). Milk urea nitrogen target concentrations for lactating dairy cows fed according to National Research Council recommendations. J. Dairy Sci., 82: 1261–1273.

Korbicz J., Obuchowicz A., Uciński P. (1994). Sztuczne sieci neuronowe – podstawy i zastosowanie. Ak. Of. Wyd. PLJ, Warszawa.

Kosiński A.R. (2002). Sztuczne sieci neuronowe. Wyd. WNT, Warszawa.

Kosiński A.R. (2014). Sztuczne sieci neuronowe – dynamika nieliniowa i chaos. Wyd. WNT, Warszawa.

Niżewska P., Boniecki P., Dach J. (2007). Ocena zastosowania prognostycznej sieci neuronowej w modelowaniu emisji gazowych. J. Res. Appl. Agr. Eng., 52 (2): 71–74.

Rajala-Schultz P.J., Saville W.J.A. (2003). Sources of variation in milk urea nitrogen in Ohio dairy herds. J. Dairy Sci., 86: 1653–1661.

Rzewulska K., Strabel T. (2013 a). Genetic parameters for milk urea concentration and milk traits in Polish Holstein-Friesian cows. J. Appl. Genet., 54: 473–482.

Rzewulska K., Strabel T., (2013 b). Effect of some non-genetic factors on concentration of urea in milk in Polish Holstein-Friesian cows. J. Anim. Feed Sci., 22: 197–203.

Satoła A., Ptak E. (2016). Wpływ wybranych czynników na zawartość mocznika w mleku krów rasy polskiej holsztyńsko-fryzyjskiej. Rocz. Nauk. PTZ, 12, 3: 21–32.

Sawa A., Bogucki M., Jankowska M., Krężel-Czopek S. (2010). Wpływ wybranych czynników na udział prób mleka o określonej zawartości białka i mocznika. Acta. Sci. Pol., 9: 57–64.

ESTIMATING THE EFFECT OF SOME MILK COMPONENTS ON UREA CONCENTRATION USING ARTIFICIAL NEURAL NETWORKS

Summary

The objective of the study was to use the advantages of artificial neural networks to determine the impact of consecutive lactations on milk urea. In the observations factors such as milk fat percentage, protein percentage, fat to milk ratio, and lactose were used. The study used data from test-day milk from a production farm in the Dolnośląskie voivodeship. The analysis revealed that the effect of selected milk components in different lactations changes based on the chosen SSN model. The obtained results may confirm the influence of next lactation on urea and selected components in milk. The use of artificial neural network in research can be a helpful instrument to model and examine nonlinear and complex correlations between factors affecting its course.

Key words: neural networks, urea, milk composition

Phot. B. Borys