

# The use of data mining techniques for analysing factors affecting cow reactivity during milking

## Wykorzystanie technik data mining w analizowaniu czynników wpływających na reaktywność krów podczas doju

Wojciech NEJA<sup>1\*</sup>, Dariusz PIWCZYŃSKI<sup>2</sup>, Sylwia KRĘŻEL-CZOPEK<sup>1</sup>, Anna SAWA<sup>1</sup> and Serkan OZKAYA<sup>3</sup>

<sup>1</sup>Department of Cattle Breeding, Faculty of Animal Science and Biology, University of Science and Technology in Bydgoszcz, Mazowiecka 28, 85-084 Bydgoszcz, Poland, \*correspondence: nejaw@utp.edu.pl

<sup>2</sup>Department of Genetics and Animal Breeding, Faculty of Animal Science and Biology, University of Science and Technology in Bydgoszcz, Mazowiecka 28, 85-084 Bydgoszcz, Poland

<sup>3</sup>Suleyman Demirel University, Agriculture Faculty, Department of Animal Science, Isparta Turkey

### Abstract

Motor activity of 158 Polish Holstein-Friesian cows was evaluated 5 times (before and during milking in a DeLaval 2\*10 milking parlour) for both the morning and evening milking, on a 5-point scale, according to the method of Budzyńska et al. (2007). The statistical analysis used multiple logistic regression and classification trees (Enterprise Miner 7.1 software which comes in with SAS package). In the evaluation of motor activity, cows that were among the first ten to enter the milking parlour were more often given a score of 3 points before (11.5%) and during milking (23.5%) compared to the other cows. Cows' activity tended to decrease (both before and during milking) with advancing lactation. The cows' reduced activity was accompanied by shorter teat cup attachment times and lower milk yields. The criteria calculated for the quality of models based on classification tree technique as well as logistic regression showed that similar variables were responsible for the reactivity of cows before milking (teat cup attachment time, day of lactation, number of lactation, side of the milking parlour) and during milking (day of lactation, side of the milking parlour, morning or evening milking, milk yield, number of lactation). At the same time, the applied methods showed that the determinants of the cow reactivity trait are highly complex. This complexity may be well explained using the classification tree technique.

**Keywords:** behavioural reactivity, data mining, logistic regression, milking parlour

## Streszczenie

Aktywność ruchową (przed i w czasie doju w hali udojowej DeLaval 2 \* 10) 158 krów phf oceniono 5-krotnie, uwzględniając każdorazowo dój ranny i wieczorny, w skali 5 pkt., według metodyki Budzyńskiej i wsp. (2007). W opracowaniu statystycznym wykorzystano wieloraką regresję logistyczną i drzewa klasyfikacyjne (oprogramowanie Enterprise Miner 7.1 wchodzące w skład pakietu SAS). Stwierdzono, że krowy, które wchodziły do hali udojowej w pierwszej dziesiątce, w ocenie aktywności ruchowej przed i podczas doju częściej (11,5% oraz 23,5%) niż w przypadku pozostałych uzyskiwały 3 pkt. Odnotowano tendencję do zmniejszenia aktywności ruchowej krów (zarówno przed jak i w czasie doju) wraz z zaawansowaniem laktacji. Wykazano, że wraz z mniejszą aktywnością ruchową krów skracał się czas zakładania kubków udojowych, jednocześnie też zmniejszała się wydajność mleka. Obliczone kryteria jakości modeli budowanych w oparciu o technikę drzew klasyfikacyjnych oraz regresji logistycznej wskazały podobne zmienne odpowiedzialne za reaktywność krów przed (czas zakładania kubków, kolejny dzień laktacji, kolejna laktacja i zajmowana strona hali udojowej) i w trakcie doju (kolejny dzień laktacji, zajmowana strona hali udojowej, dój ranny lub wieczorny, wydajność mleka oraz kolejna laktacja). Jednocześnie zastosowane metody wskazały znaczną złożoność uwarunkowania cechy, jak jest reaktywność krów. Złożoność ta może być dobrze wyjaśniona za pomocą techniki drzew klasyfikacyjnych.

**Słowa kluczowe:** aktywność behawioralna, hala udojowa, regresja logistyczna

## Introduction

In recent years, animal scientists have shown a considerable interest in cattle behaviour (Philips, 2002; Kowalski, 2005; Budzyńska et al., 2007; Adamczyk and Gil, 2014; Neja et al., 2015). Behaviour can be defined as the animal body's overall response to environmental factors, one that is genetically determined but also modifiable as a result of adaptive processes; therefore, behavioural analysis can serve as a method to determine welfare levels in animals (Kowalski, 2005).

Milking is one of the most complex operations in the farming and production of dairy cattle. The use of milking parlours in free-stall systems interferes with the behaviour of animals and requires the cows to develop new behavioural reactions desired by handlers during milking (Paranhos da Costa and Broom, 2001).

Statistical analysis of behavioural traits in cattle can be performed using different methods and models (Budzyńska et al., 2007; Adamczyk et al., 2011; Adamczyk and Gill, 2014). The recent advances in statistical methods have led to development of new tools for data analysis, known as data mining (Lasek, 2007). Data mining techniques, which include cluster analysis, decision trees, and artificial neural networks, find application in different areas of human activity, such as finances, medicine, and animal breeding (Abu-Hanna and De Keizer, 2003; Feldman and Gross, 2003; Sawa et al., 2004; Austin, 2007; Lasek, 2007; Piwczyński and Sitkowska, 2012; Grochowska et al., 2014). One of the data mining techniques are

decision trees, where a tree model is constructed by segmenting the whole group of observations into subgroups (so-called leaves) that are maximally homogeneous in terms of dependent variable (Lasek, 2007). These divisions produce a graphic model with a tree-like structure. It consists of nodes, branches, and leaves.

The aim of the study was to perform statistical analysis of dairy cow reactivity before and during milking in a milking parlour, using classification trees and logistic regression.

## Materials and methods

The study was carried out in a free-stall barn with a DeLaval 2×10 milking parlour. Observations were made on 158 Polish Holstein-Friesian cows divided into 4 technological groups. Motor activity of each animal was recorded 5 times for both the morning and evening milking, according to the method reported by Budzyńska et al. (2007). Continuous direct observations were made for each cow between entering and leaving the milking stall, distinguishing between cow's activity before milking (during udder cleaning and teat cup attachment) and during milking. A 5-point scale arranged in decreasing order of excitability was used:

- 1 – persistent and energetic stepping and kicking, excitable cow,
- 2 – persistent and energetic stepping, no kicking,
- 3 – energetic leg movements, intermittent,
- 4 – standing with occasional, slight leg movements,
- 5 – standing, no movement of legs, very calm cow.

Cows' activity before and during milking was determined by the following factors in the statistical modelling:

- lactation number (1, 2, 3, 4 – 7),
- order of entry into the milking parlour (first, second, third, fourth, fifth) of 10 cows from different technological groups (1, 2, 3, 4, 5),
- side occupied – side occupied by a cow in the milking parlour (left (L) and right (R), within the area of stalls 1 – 3, stalls 4 – 7 and stalls 8 – 10) -L1-3, L4-7, L8-10, R1-3, R4-7, R8-10,
- time of day when the cows were milked (morning and evening milking).

In the statistical modeling cows' activity before and after milking was also determined by milking factors such as: lactation day, milk yield per milking (kg), teat cup attachment time (TCAT) (s).

Statistical analysis of the effect of these factors on the behaviour of cows before and during milking was performed in parallel classification trees and by multiple logistic regression. To this end, SAS Enterprise Miner 7.1 software was used (SAS Institute Inc. 2014).

For the statistical analysis, the group of observations divided into two sets: a training set (60%) and a validation set (40%). These sets were generated by the stratified

sampling method. Preliminary fitting of the model was the purpose of the training set. On the other hand, the validation set allowed to compare the quality of the generated models, but also to avoid overfitting (SAS Institute Inc., 2014). When building the classification tree, the minimum terminal node size was set to 30, with maximum depth of 6. It was also assumed that a binary tree will be created. When determining the leaf size and tree depth criteria, authors wanted to avoid overfitting the tree to the training data, which could result in the reflection of random relationships in the validation set (Piwczyński et al., 2013).

When constructing the classification tree, was used the Gini index as the division criterion (Piwczyński et al., 2013):

Gini index ( $G(p)$ ):

$$G(p) = 1 - \sum_{j=1}^k p_j^2$$
 where:  $p$  – the probability vector of object assignment to classes in the form of:  $p = (p_1, p_2, \dots, p_k) = \left(\frac{l_1}{n}, \frac{l_2}{n}, \dots, \frac{l_k}{n}\right)$ ,  $k$  – the number of classes,  $l$  – the size of class,  $n$  – the size of analysed population (cows).

The applied criterion of tree splitting was a result of using the CART algorithm. The developed models of classification trees belong to the Single Tree Models (SAS Institute Inc., 2014).

The Gini coefficient has a property whereby this measure takes on the value of zero when a trait distribution is centred on a single value. The higher the values they take, the more diverse a population is in terms of the examined trait.

The ranking of variables in terms of their importance in creating data set splits was prepared based on the “Importance” measure (Piwczyński et al., 2013; SAS Institute Inc., 2014).

The quality of fit of the constructed models to the data was compared using average squared error, misclassification rate, cumulative lift, Kolmogorov–Smirnov statistics and the area under the Receiver Operating Characteristic curve (ROC index) (SAS Institute Inc., 2014). Decreasing values of the first two statistics are indicative of the increasing value of the model. An inverse relationship holds for the other model quality criteria, namely a better model has higher statistic values.

The following information was added to each created node and a resulting leaf: number of node ID (1), percentage of cows with reactivity up to 3 (6.43%) (2), percentage of cows with reactivity up to 4 (41.94%) (3), percentage of cows with reactivity up to 5 (51.63%) (4), number of observations in a node or leaf (949) (5).

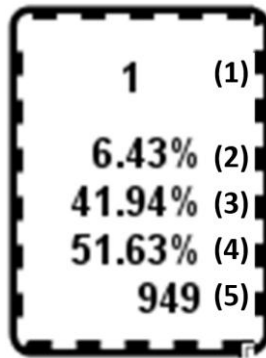


Figure 1. Description of a node

Rycina 1. Opis węzła

For the purpose of comparison, statistical analysis using multiple logistic regression was also performed. When constructing the regression model, stepwise forward selection of variables was used (SAS Institute Inc., 2014), and 1 interactions were accounted for in addition to factors (variables).

## Results and discussion

Table 1 shows the percentage of cows with different reactivity before and during milking, depending on selected factors. The observed dairy cows were less active before milking (only 6% of the cows made energetic leg movements (3 pts.) and as much as 52% were assessed as calm (5 pts.) than during milking (as much as 15% with a score of 3 pts. and almost 40% with a score of 5 pts). The results obtained are in agreement with the findings of Budzyńska et al. (2007). With the age of the cows (consecutive lactation), the proportion of animals whose activity was scored 3 pts. tended to decline, with a decrease from 7.3% to 3.5% before milking, and from 17.2% to 7.0% during milking. The percentage of excitable cows was observed to be higher during morning than evening milking. In the consecutive tens of cows that entered the milking parlour, the proportion of excitable cows (3 pts.) decreased from 11.5% to 2.1% (reactivity before milking) and from 23.5% to 10.5% (reactivity during milking). Budzyńska et al. (2007) reported that more excitable cows were the first to occupy the stalls in the milking parlour. Regardless of the side of the milking parlour and the area occupied, the proportion of restless cows (3 pts.) was mostly over twice as low before milking than during milking. Paranchos da Costa and Broom (2001) found no evidence that dairy cows were discomforted when milked in the non-preferred side of the milking parlour.

Table 1. Cows' reactivity before and during milking depending on the experimental factors

Tabela 1. Reaktywność krów przed i w czasie doju w zależności od badanych czynników

Factor	Level	Statistic	Reactivity before milking (pts)				Reactivity during milking (pts)			
			3	4	5	Total	3	4	5	Total
No. of lactation	1	n	63	388	409	860	148	379	332	859
		%	7.33	45.12	47.56		17.23	44.12	38.65	
	2	n	16	113	161	290	42	130	118	290
		%	5.52	38.97	55.52		14.48	44.83	40.69	
	3	n	15	91	124	230	33	114	83	230
		%	6.52	39.57	53.91		14.35	49.57	36.09	
	4 - 7	n	7	71	122	200	14	94	92	200
		%	3.5	35.5	61		7	47	46	
Milking time	morning	n	54	375	361	790	129	377	283	789
		%	6.84	47.47	45.7		16.35	47.78	35.87	
	evening	n	47	288	455	790	108	340	342	790
		%	5.95	36.46	57.59		13.67	43.04	43.29	
Order of entry into the milking parlour	1	n	46	189	165	400	94	173	133	400
		%	11.5	47.25	41.25		23.5	43.25	33.25	
	2	n	31	163	205	399	58	174	167	399
		%	7.77	40.85	51.38		14.54	43.61	41.85	
	3	n	16	142	183	341	47	170	124	341
		%	4.69	41.64	53.67		13.78	49.85	36.36	
	4	n	3	83	114	200	13	88	99	200
		%	1.5	41.5	57		6.5	44	49.5	
	5 - 6	n	5	86	149	240	25	112	102	239
		%	2.08	35.83	62.08		10.46	46.86	42.68	
Side occupied by a cow in the milking parlour (L – left side R – right side)	L 13	n	13	115	115	243	38	128	77	243
		%	5.35	47.33	47.33		15.64	52.67	31.69	
	L 47	n	23	142	132	297	31	148	118	297
		%	7.74	47.81	44.44		10.44	49.83	39.73	
	L 810	n	9	106	101	216	20	108	88	216
		%	4.17	49.07	46.76		9.26	50	40.74	
	P 13	n	24	90	153	267	62	102	103	267
		%	8.99	33.71	57.3		23.22	38.2	38.58	
	P 47	n	20	122	181	323	57	132	133	322
		%	6.19	37.77	56.04		17.7	40.99	41.3	
P 810	n	12	88	134	234	29	99	106	234	
	%	5.13	37.61	57.26		12.39	42.31	45.3		
Total			101	663	816	1580	237	717	625	1579

Analysis of the results in Table 2 shows that milk yield (12.7 kg) was lowest in very calm cows (5 pts.) and increased with increasing reactivity of the cows (both before and during milking). Teat cup attachment time increased with increasing motor activity of the cows (by about 1 s). The cows' reactivity decreased as lactation advanced. Most of the results of other authors also demonstrate that the cows' milking behaviour does affect their milk yield. According to Maffei et al. (2006), especially in Holstein-Friesian cows milk yield is higher for nervous and very nervous cows than for calm cows. Praxedes et al. (2011) reported that the higher handling costs for nervous cows are compensated by their higher milk yield.

Table 2. Behaviour of the cows depending on the experimental factors

Tabela 2. Zachowanie się krów w zależności od badanych czynników

Variable	Statistic	Reactivity before milking			Reactivity during milking			Total
		(pts.)			(pts.)			
		3	4	5	3	4	5	
	N	101	663	815	237	717	625	1579
Milk yield (kg)	Mean	15.2	13.4	12.71	15.04	12.96	12.68	13.16
	Std	5.18	5.19	5.42	5.59	5.11	5.37	5.35
Teat cup attachment time (s)	Mean	6.66	5.99	5.61	6.26	5.83	5.69	5.84
	Std	1.62	1.21	1.36	1.59	1.28	1.29	1.35
Days of lactation (days)	Mean	135.81	183	206.87	157.56	202	194.36	192.31
	Std	128.21	131.5	135.84	144.77	137.38	125.62	134.77

The measures used to assess the quality of the classification tree model and of logistic regression show that both statistical methods have similar predictive ability (Table 3). At the same time they allow a conclusion that statistical models for describing the reactivity of animals prior to milking have a better quality than those during milking, as evidenced by the calculated measures of model quality (Table 3). All the ROC indices assumed relatively low values, which suggests that both logistic regression and classification tree technique have limited predictive value for cow reactivity before and during milking (Lasek, 2007). Higher values of the area above the ROC curve (0.611-0.711) were reported in the present author's earlier studies (Piwczyński and Sitkowska, 2012; Piwczyński et al., 2013).

Table 3. Model comparisons  
Tabela 3. Porównanie modeli

Statistic label	Reactivity before milking		Reactivity during milking	
	Logistic regression	Classification trees	Logistic regression	Classification trees
Kolmogorov-Smirnow statistic	0.224	0.171	0.137	0.112
Average Squared Error	0.1766	0.1758	0.2071	0.2034
ROC index	0.642	0.608	0.572	0.569
Cumulative lift	1.3912	1.561	1.106	1.5307
Misclassification rate	0.4374	0.458	0.5491	0.5316

Table 4 presents the importance of different variables in creating a graphic model of the classification tree describing the reactivity of cows before and during milking. The results obtained prove that cow reactivity before milking was influenced most by the teat cup attachment time, day of lactation, order of entry, and the place occupied in the parlour. In turn, the reactivity during milking was diversified most by day of lactation, place occupied in the parlour, time of day, and, to a similar extent, by milk yield and number of lactation.



Table 4. Variables importance (GINI)  
Tabela 4. Znaczenie zmiennych (GINI)

Variable	Reactivity before milking	Reactivity during milking
Teat cup attachment time	1	0
Days of lactation	0.6805	1
Milking time	0	0.5105
Order of entry into the milking parlour	0.6323	0
No. of lactation	0	0.4067
Milk yield	0	0.4116
Side occupied by a cow in the milking parlour	0.559	0.5826

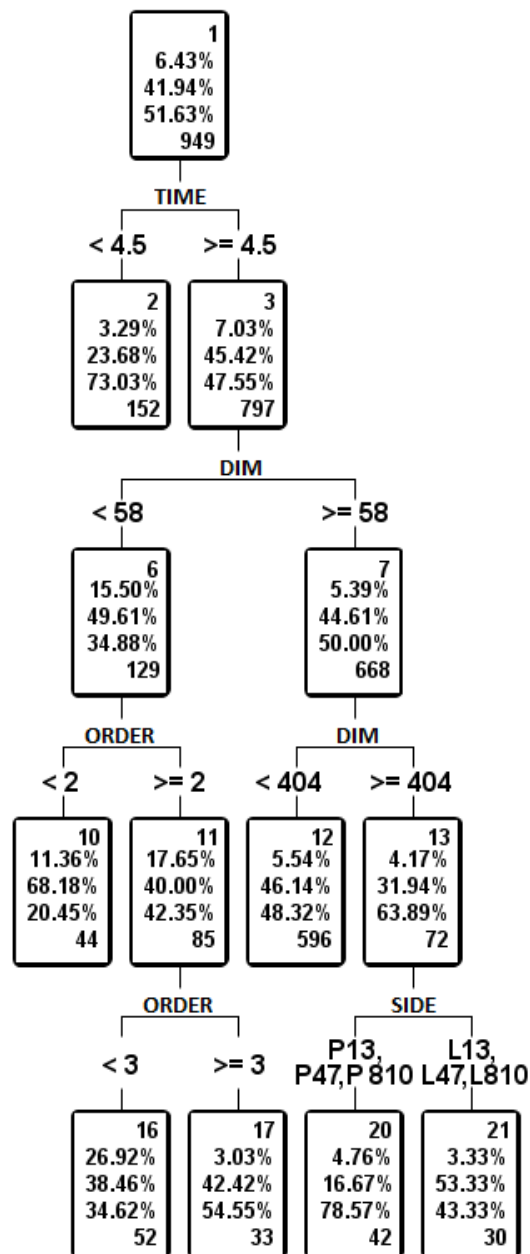
Logistic regression analysis of factors that affect cow behaviour (Table 5) revealed that they include teat cup attachment time, order of entry into the milking parlour, place occupied in the milking parlour, day of lactation, and time of day for reactivity before milking; and milk yield, teat cup attachment time, order of entry into the milking parlour, time of day, and place occupied in the milking parlour for reactivity during milking.

Table 5. Variables responsible for the behaviour of cows determined by logistic regression

Tabela 5. Zmienne odpowiedzialne za zachowanie się krów wskazane przy użyciu regresji logistycznej

Variable	Wald Chi- Square Test	P
Reactivity before milking		
Teat cup attachment time	34.9459	<0.0001
Milking time	4.1477	0.0417
Order of entry into the milking parlour	20.0631	0.0005
Days of lactation	15.2863	<0.0001
Side occupied by a cow in the milking parlour	18.8793	0.002
Milking time * Side occupied by a cow in the milking parlour	15.7333	0.0076
Reactivity during milking		
Teat cup attachment time	12.1323	0.0005
Order of entry into the milking parlour	14.8622	0.005
Milk yield	19.6851	<0.0001
Side occupied by a cow in the milking parlour	11.1418	0.0486
Milking time * Side occupied by a cow in the milking parlour	11.7985	0.0377
Order of entry into the milking parlour × no. of lactation	21.8318	0.0394

The classification tree for cows' activity before milking (Figure 2), based on Gini coefficient reduction, has 6 leaves and is 4 levels deep. The most important factor differentiating the cows' activity was the teat cup attachment time, on the basis of which two sets were identified: <4.5 s (node 2) and ≥4.5 s (node 3). The former set was characterized by an approximately twice lower proportion of cows whose activity was scored as 3 and 4 pts. In a study by Budzyńska et al. (2007), the mean teat cup attachment time of 7.84 s was positively correlated with reactivity during milking and negatively with reactivity before milking.



Abbreviations: TIME - teat cup attachment time (s), ORDER - order of entry into the milking parlour (first, second, third, fourth, fifth) of 10 cows from different technological groups (1, 2, 3, 4, 5), DIM - days of lactation (days), SIDE - side occupied by a cow in the milking parlour (L – left side of the milking parlour, R – right side of the milking parlour).

Figure 2. Splits of node 1 for cows' characterizing reactivity before milking

Wykres 2. Podziały węzła 1 charakteryzującego reaktywność krów przed dojem

The first set of the cows was not further partitioned and became a leaf. Node 3 was partitioned based on the milking day variable and resulted in a subset of cows up to 58 days of lactation (node 6) and those beyond 58 days of lactation (node 7). Among the cows up to 58 days of lactation, the proportion of cows with highest motor activity

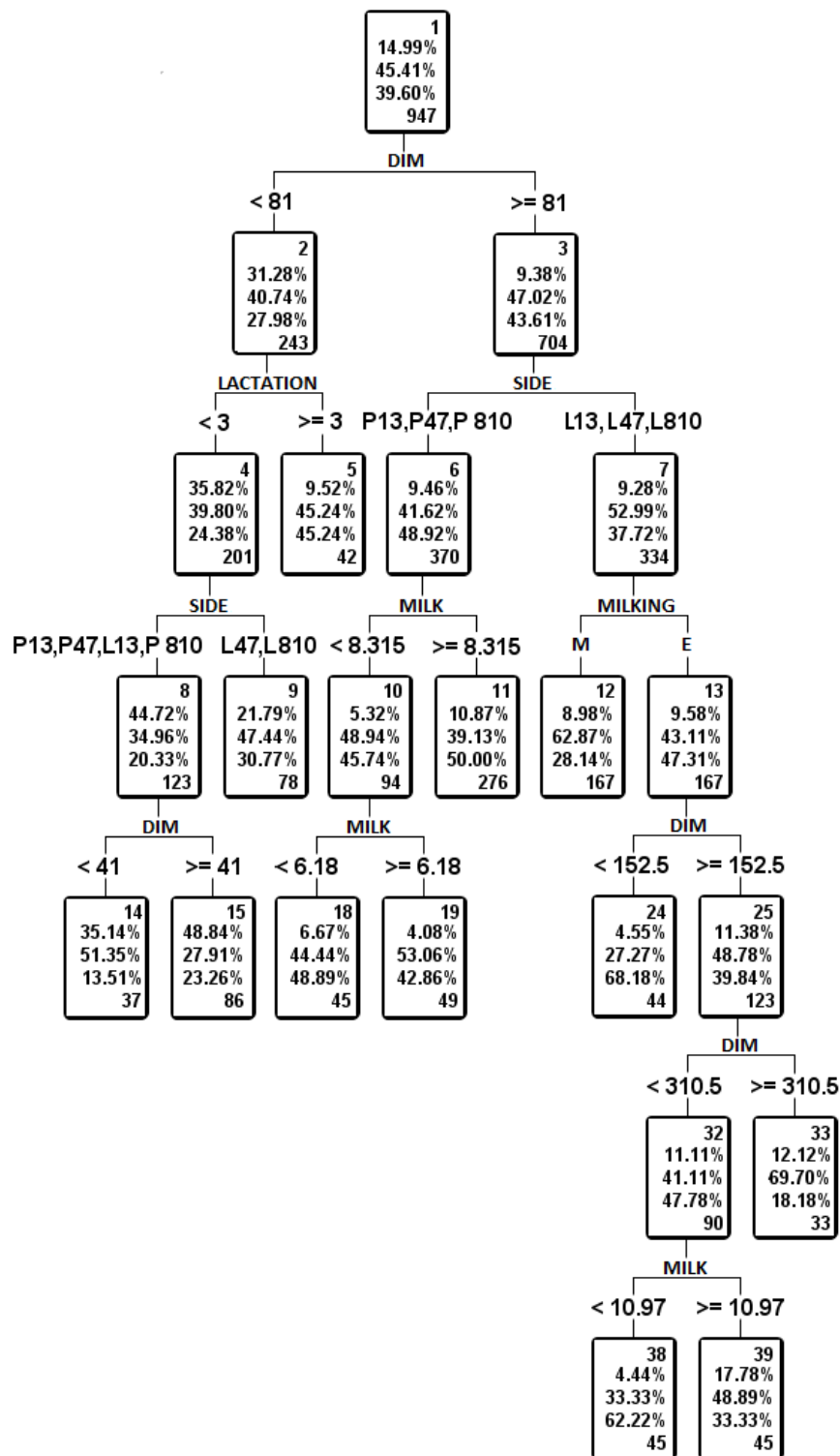
(3 pts.) was 3 times as high as for cows at a more advanced stage of lactation, which confirms the results presented in Table 2.

The factor that differentiated the early lactation cows (node 6) was the order of entry into the milking parlour: the first twenty (subset 10) and the remaining cows (subset 11). In the first subset (10), which was not further partitioned, 68% of the cows were scored as 4 pts. and only 20% as 5 pts. In node 11 there were over 42% of very calm cows (5 pts.). According to Budzyńska et al. (2007), more excitable cows were the first to occupy a stall in the milking parlour. Earlier research by one of the present authors showed that highest yielding cows more often entered the parlour as one of the first ten or twenty cows compared to others (Neja et al., 2006). According to Galindo and Broom (2000), the order of entry into the milking parlour may be determined by actual daily milk yield, stage of reproductive cycle, and health status of the cows.

In this case, too, it was found that cows that were more reluctant to enter the milking parlour (node 17) were more often calm before milking. The largest proportion (almost 27%) of restless cows (3 pts.) was found in node 16 for cows that were among the first thirty to enter the parlour.

The factor differentiating the cows from day 58 of lactation (node 7) was again the day of lactation (subset of cows up to 404 days of lactation (node 12) and beyond 404 days of lactation (node 13). As much as 64% of the cows from the last subset were characterized by low reactivity (5 pts.). This node was further partitioned based on the place occupied by a cow in the milking parlour. Almost 80% of the cows on the right side of the milking parlour (node 20) were assessed to be very calm (5 pts.). According to Dobicki et al. (2002) the tendency for regular choice of the left or right side of the milking parlour by the cows is similar to normal distribution. The authors showed high repeatability for the stall number chosen by a cow ( $0.25 \leq r \leq 0.64$ ). At the same time, they found significant repeatability for the order of cows entering the milking parlour in the case of all milkings, and the calculated coefficient of correlation ( $=0.546$ ) may indicate that this trait was a permanent herd behaviour.

The model of the decision tree for cow reactivity during milking was more extended 12 (12 leaves and 6 levels) (Figure 3). The first partition was based on the lactation day variable, giving node 2 (<81 days) and node 3 ( $\geq 81$  days). Among the cows up to day 81 of lactation, as much as 31% made energetic intermittent leg movements (3 pts.), 41% were standing with occasional, slight leg movements (4 pts.), and 28% stood with no movement of legs (5 pts.). Among the cows in later stages of lactation, the proportion of most excitable cows was 3 times as low, while the proportion of calmer cows increased (4 and 5 pts.). This result confirms the trends observed for cows' premilking activity, namely that the proportion of calm cows increases as lactation progresses. Both subsets were further partitioned, the former by the lactation number (<3 and  $\geq 3$ ), and the latter by the side occupied in the parlour (left and right).



Abbreviations: DIM - days of lactation (days), SIDE - side occupied by a cow in the milking parlour (L – left side of the milking parlour, R – right side of the milking parlour), LACTATION - number of lactation (1, 2, 3, 4-7), MILK - milk yield per milking (kg), MILKING - time of day when the cows were milked (morning and evening milking).

Figure 3. Splits of node 1 for cows' characterizing reactivity during milking  
Wykres 3. Podziały węzła 1 charakteryzującego reaktywność krów w czasie doju

Cows up to the third lactation (node 4) were characterized by higher activity during milking compared to older cows (almost 36% with a score of 3 pts. and only 24% with a score of 5 pts.). The subset of older cows was not further partitioned (it became a leaf) and came to include mostly cows standing with occasional, slight leg movements (45%) and those standing motionless (45%). Node 4 was partitioned based on the place occupied in the milking parlour. Almost 45% of cows occupying the right side of the parlour as well as stalls 1-3 on the left side of the parlour (node 8) were classified as excitable, because they made energetic leg movements during milking (3 pts.). Cows occupying the other place in the milking parlour (node 9) formed a subset (leaf), in which the excitable cows did not exceed 22%, forming the smallest group. Node 8 was partitioned based on the lactation day into subsets of <41 and  $\geq$ 41 days. The peak lactation cows (node 15) were dominated (49%) by excitable cows that made energetic leg movements during milking (3 pts), while there were only 23% of very calm cows.

It was found that the most important factor that differentiated the cows beyond 81 days of lactation (node 3) was the place occupied in the milking parlour, but the differences concerned the proportion of motor activity scored as 4 and 5 pts. The node for cows occupying the right side of the milking parlour (node 6) was partitioned based on milk yield per milking, and again the milk yield was the factor that differentiated the subset of lower yielding cows. In the evaluation of motor activity, the lower yielding cows were more often assessed to be calm (5 pts.). Analysis of the group of cows occupying the left side of the parlour (node 7) showed almost 53% of the cows scoring 4 pts. for activity and 37% of very calm cows. Node 7 was partitioned based on the time of milking: node 12 (morning) with 63% of calm and 28% of very calm cows, and node 13 (evening) with over 43% of calm and 47% of very calm cows. Node 12 was not partitioned further, whereas node 13 was differentiated based on milking days (cows up to 152 days of lactation (node 24) and cows beyond 152 days of lactation (node 25), with node 24 becoming a leaf and node 25 being partitioned again based on milking day. The proportion of cows which scored 5 pts. for activity was found to decrease with advancing lactation ( $\geq$ 152 days and  $\geq$ 310 days). Node 33 was not further partitioned, whereas node 32 was split, based on milk yield: <10.97 kg milk (node 38) and  $\geq$ 10.97 kg (node 39), in which the most restless cows formed as much as 17.7%. In the group of lower yielding cows (node 38), 62% of the cows were very calm.

## Conclusions

In the evaluation of motor activity, it was found that more excitable cows were the first to occupy the stalls in the milking parlour, their milk yields was greater.

Teat cup attachment time increased with increasing motor activity of the cows (by about 1 s). The cows' reactivity decreased as lactation advanced.

The criteria calculated for the quality of models based on classification tree technique as well as logistic regression showed that similar variables were responsible for the reactivity of cows before milking (teat cup attachment time, day of lactation, number of lactation, side of the milking parlour) and during milking (day of lactation, side of the milking parlour, morning or evening milking, milk yield, number of lactation). At

the same time, the applied methods showed that the determinants of the cow reactivity trait are highly complex.

## References

- Abu-Hanna, A., De Keizer, N. (2003) Integrating classification trees with local logistic regression in intensive care prognosis. *Artificial Intelligence in Medicine*, 29, 5–23. DOI: [https://doi.org/10.1016/S0933-3657\(03\)00047-2](https://doi.org/10.1016/S0933-3657(03)00047-2)
- Adamczyk, K., Słania, A., Gil, Z., Felenczak, A., Bulla, J. (2011) Relationships between milk performance and behavior of cows under loose housing conditions. *Annals of Animal Science*, 11 (2), 283–293.
- Adamczyk, K., Gil, Z. (2014) Evaluation of welfare of dairy cattle based on behavioral indicators. *Zootechnical News*, 52 (2), 43–49.
- Austin, P. C. (2007) A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AML mortality. *Statistics in Medicine*, 26, 2937–2957.
- Budzyńska, M., Krupa, W., Tietze, M. (2007) Behavior of cows in the milking parlor. *Veterinary Medicine*, 63 (11), 1363–1365.
- Dobicki, A., Chudoba, K., Nowopolska, A., Piestrak, S. (2002) A herd of cows behavior during the occupation of posts in the milking parlor. In: 7th International Scientific Conference: Theoretical and applicatory problems and agricultural engineering in the process adaptation to EU research programs. Polanica Zdrój, Poland, 2002, 475–478.
- Feldman, D., Gross, S. (2003) Mortgage default: classification trees analysis. The Pinhas Sapir Center for Development Tel-Aviv University, Discussion Paper, 3, 1–46.
- Galindo, F., Broom, D. M. (2000) The relationship between social behavior of dairy cows and the occurrence of lameness in three herds. *Research in Veterinary Science*, 69 (1), 75–79. DOI: <https://doi.org/10.1053/rvsc.2000.0391>
- Grochowska, E., Piwczyński, D., Portolano, B., Mroczkowski, S. (2014) Analysis of the influence of the PrP genotype on the litter size in Polish sheep using classification trees and logistic regression. *Livestock Science*, 159, 11–17. DOI: <http://dx.doi.org/10.1016/j.livsci.2013.11.008>
- Kowalski, A. (2005) Stereotypies as an indicator of animal welfare. *Veterinary Medicine*, 61, 1335–1339.
- Lasek, M. (2007) Methods of Data Mining in analyzing and forecasting of economic condition of enterprises. *Application Enterprise Miner*, Difin Warsaw, 1–174.
- Maffei, W., Bergmann, J., Pinotti, M., Lima Neto, H. (2006) Reactivity in a mobile cage - a new methodology to evaluate bovine temperament. In: 8th World Congress on Genetics Applied to Livestock Production. Belo Horizonte, Minas Gerais, Brazil, 13–18 August 2006.

- Neja, W., Sawa, A., Kiedrowicz, M. (2006) The behavior of the cows during the occupation of posts in the milking parlor. *Scientific Annals of Polish Society of Animal Production*, 2 (2), 127–132.
- Neja, W., Sawa, A., Jankowska, M., Bogucki, M., Krężel-Czopek, S. (2015) Effect of the temperament of dairy cows on lifetime production efficiency. *Archives Animal Breeding*, 58, 193–197.
- Paranhos da Costa, M. J. M., Broom, D. M. (2001) Consistency of side choice in the milking parlour by Holstein-Friesian cows and its relationship with their reactivity and milk yield. *Applied Animal Behaviour Science*, 70, 177–186.
- Philips, C. (2002) *Cattle behavior and welfare*. Blackwell Sciences, 14.
- Piwczyński, D., Sitkowska, B. (2012) Statistical modelling of somatic cell counts using the classification tree technique. *Archives Animal Breeding*, 4 (55), 332–345.
- Piwczyński, D., Nogalski, Z., Sitkowska, B. (2013) Statistical modeling of calving ease and stillbirths in dairy cattle using the classification tree technique. *Livestock Science*, 154, 19–27. DOI: <https://doi.org/10.1016/j.livsci.2013.02.013>
- Praxedes, V., Verneque, R., Pereira, M., Pires, M., Machado, M., Peixoto, M. (2011) Evaluation of factors influencing milk ejection in the Brazilian Gyr dairy cattle. *Brazilian Agricultural Research Corporation (EMBRAPA)*.
- SAS Institute Inc. (2014) *SAS/STAT 9.4 User's Guide*. Cary, NC: SAS Institute Inc.
- Sawa, A., Kowaliszyn, B., Piwczyński, D. (2004) The application of ID3 algorithm in the selection of the genetic and physiological parameters-the environmental conditions obtaining milk with low somatic cell count and high protein content. *Scientific Annals of Polish Society of Animal Production*, 72 (1), 41–49.