lekarzami weterynarii i korektorami racic jest bardzo cenna. Zamierzamy w jak najkrótszym czasie uzyskać najlepsze efekty, dlatego nie zakładamy, że stworzymy jednolity system zbierania informacji, który obejmie cały kraj, tak jak to jest za granicą. Ale nie we wszystkich krajach takie systemy funkcjonują, a mimo to ocena cech zdrowia jest prowadzona. Zatem każdy sposób zbierania informacji jest dobry – mogą to być dane od lekarzy weterynarii lub od hodowcy. Źródłem informacji mogą być także programy do zarządzania stadem. Trzeba wyraźnie pokreślić, że bez zaangażowania się hodowców w zbieranie danych, np. o chorobach wymion, Centrum nie będzie mogło im pomóc. Mało tego, na korzyści z tej współpracy hodowca będzie musiał czekać rok, dwa lata. Na szczęście dzięki ocenie genomowej dane, np. o mastitis, można zbierać z grupy krów, a nie z całej populacji objętej kontrolą użytkowości.

Czy prowadzenie selekcji bydła na kilkadziesiąt cech jest efektywne?

Selekcja genomowa przyspieszyła rotację pokoleń. Dzięki temu bardziej efektywna stała się selekcja na dużą liczbę cech, w tym także nisko odziedziczalnych. Ponadto ocena genomowa zwiększyła dokładność doskonalenia cech o niskiej odziedziczalności, np. zdrowia i płodności. W efekcie wagi tych cech w indeksach selekcyjnych przyjmują większe wartości.

Teraz wyniki oceny wartości użytkowej buhajów publikowane są w Polsce 3 razy w roku, w innych krajach co miesiąc. Jaka jest korzyść tak częstej oceny wartości hodowlanej rozpłodników?

W niektórych krajach takie oceny przeprowadzane są co tydzień. Im częściej wykonywana jest ocena, tym większa korzyść dla hodowcy, ponieważ szybciej uzyskuje on aktualną informację o wartości hodowlanej zwierząt. Jeżeli na taką informację czeka pół roku, to oznacza, że przez ten czas nie wie, jaka jest wartość hodowlana zwierzęcia i nie może podjąć trafnej decyzji, czy buhajka wybrakować, czy zostawić do dalszej hodowli. Przez ten czas zwierzę jest żywione, co wiąże się z kosztami. Informacja opóźniona jest zawsze mniej użyteczna. Wartość hodowlana samic powinna być także możliwie jak najczęściej aktualizowana. Oprócz młodzieży warto genotypować starsze osobniki, aby mieć informację o ich aktualnej wartości hodowlanej. Jest to szczególnie przydatne np. przy wyborze krów na matki buhajów. Co więcej, pomaga oddzielić samice, które powinny stać się matkami przyszłych córek, od tych, które można kryć nasieniem buhajów ras mięsnych lub przeznaczyć na biorczynie. Wszystko po to, by zwiększyć zyski.

Czy obecnie są bariery natury organizacyjnej ograniczające pełne wykorzystanie potencjału nowo powstałego Centrum Genetycznego?

Tworząc Centrum wiedzieliśmy, jakie są uwarunkowania w krajowej hodowli bydła mlecznego i widzimy ogromny sens, aby w tych warunkach funkcjonować. W Polsce doszło do wyjątkowej sytuacji, w której hodowcy z własnej inicjatywy założyli takie Centrum i sami je finansują. Byłoby ono o wiele lepiej wykorzystane, gdyby wykonywało rutynowo ocenę wartości hodowlanej. Obecnie zajmuje się tym Instytut Zootechniki. Moim zdaniem, gdyby skład osobowy Centrum poszerzyć o naukowców zaangażowanych obecnie w ocenę wartości hodowlanej, to uzyskalibyśmy potencjał porównywalny z tym, którym dysponują najlepsze centra obliczeniowe za granicą. Problemem jest przekonanie Ministerstwa Rolnictwa do zmodernizowania oceny wartości hodowlanej, a to się nie stanie, jeśli nie będzie ona przekazana hodowcom. Instytut Zootechniki przez ostatnie 1,5 roku nie przeprowadził oceny genomowej z wykorzystaniem pełnej bazy EuroGenomics. To jest ostatni sygnał do przeprowadzenia zmian.

Dziękuję za rozmowę (rozmawiał **Ryszard Lesiakowski).**

Modelowanie przebiegu porodu przy użyciu drzew klasyfikacyjnych

Marcin Brzozowski, Beata Sitkowska, Dariusz Piwczyński

Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy

Intensywne doskonalenie genetyczne cech produkcyjnych bydła pozwoliło na uzyskanie zwierząt o wydajnościach życiowych przekraczających 100 tys. kg mleka oraz stad, w których przeciętna wydajność wynosi ponad 14 tys. kg mleka od krowy [17]. Jednocześnie coraz częściej hodowcy wysokowydajnych krów wskazują na problemy z szeroko rozumianym rozrodem swoich zwierząt. Wielu autorów wskazuje płodność krów jako kluczowy czynnik odpowiedzialny za opłacalność produkcji mleka w stadach bydła [5, 7, 8, 22]. Jak podkreśla Banos [1], poród jest końcowym etapem ciąży, a jego przebieg determinują następujące czynniki: warunki utrzymania, genotyp i wiek matki oraz płeć cielęcia. W stadach wysokowydajnych zaobserwowano zależności między wskaźnikami płodności krów a ich wydajnością mleczną. Krzyżewski i Reklewski [9] podkreślają negatywny wpływ wysokiej wydajności na płodność zwierząt. Niska skuteczność zacieleń, brak objawów rui czy szybkie jałowienie są konsekwencjami tej negatywnej zależności. W badaniach Bicalho i wsp. [4] konsekwencjami skomplikowanych porodów były spadek wydajności mlecznej i płodności u krów oraz wzrost kosztów związanych z opieką weterynaryjną. Wyniki badań De Vries [5] dowodzą, że martwe urodzenia i problemy z porodem wpływają znacząco na koszty bezpośrednie (utrata cielęcia, śmierć krowy) i pośrednie (pogorszona płodność, obniżona wydajność mleczna, koszty związane z koniecznością leczenia zwierząt). Z tego też względu w ostatnim czasie uwaga hodowców koncentruje się na doskonaleniu cech funkcjonalnych, do których należą cechy płodności. Mee i wsp. [12] podają, że frekwencja problemów z rozrodem u różnych ras bydła (holsztyńsko-fryzyjskie, norweskie czerwone, szwedzkie czerwone i szwedzkie białe) w różnych krajach europejskich, USA, Australii i Nowej Zelandii kształtuje się w szerokim przedziale 3-23%. Problemy z rozrodem krów przyczyniają się do krótkiego okresu użytkowania i konieczności wczesnego brakowania zwierząt [3]. W Danii, ze względu na znaczenie ekonomiczne, cechy płodności (łatwość wycieleń i poronienia) włączono do programu hodowlanego bydła mlecznego już ponad dziesięć lat temu [8].

Liczne badania naukowe [2, 6, 14, 20] dowodzą, że możliwe jest zastosowanie technik *data mining* do prognozowania pewnych zmian czy procesów związanych ze zwierzętami. Przykładami takich zastosowań są m.in. badania prowadzone przez różnych autorów nad cechami związanymi z rozrodem, tj. śmiertelnością jagniąt [19] oraz cieląt [2, 6, 14, 23]. Grzesiak i wsp. [6] na podstawie przeprowadzonych badań stwierdzili, że zastosowanie algorytmu CART może wesprzeć rolników w poprawie płodności krów w stadach, przez bezpośrednie wskazanie grupy zwierząt o prognozowanej niższej skuteczności za-

cieleń. Zaborski i wsp. [24] na podstawie przeprowadzonych analiz za pomocą metod *data mining* stworzyli model dobrej jakości, który stosunkowo dobrze wykrywał zwierzęta z dystocją.

W prezentowanym opracowaniu modelowano statystycznie przebieg porodu pierwiastek i wieloródek rasy polskiej holsztyńsko-fryzyjskiej za pomocą jednej z technik *data mining*, tj. drzew klasyfikacyjnych. Zastosowany rodzaj analizy statystycznej wynikał z jakościowego charakteru cechy, a zatem wykluczał użycie testów parametrycznych.

Materiał zwierzęcy stanowiły krowy rasy polskiej holsztyńsko-fryzyjskiej pochodzące z 7 wysokowydajnych stad zlokalizowanych na terenie całej Polski. Analizowane dane dotyczyły 23 555 krów, wśród których 9525 stanowiły pierwiastki, zaś pozostałe 14 030 szt. – wieloródki. W większości badanych stad liczba zwierząt znajdowała się na podobnym poziomie. Wyjątek stanowiło stado D (składające się z czterech obór), w którym znajdowało się około 50% wszystkich badanych sztuk. We wszystkich analizowanych stadach panowały zbliżone warunki środowiskowe.

Badany materiał zwierzęcy kontrolowano w zakresie przebiegu porodu, który został zakodowany w postaci czteropunktowej skali: 1 – poród samodzielny, 2 – poród łatwy, 3 – poród trudny, 4 – poród ciężki lub cesarskie cięcie.

Przebieg porodu w trakcie modelowania statystycznego warunkowano (tab. 1) następującymi czynnikami: stado (A, ..., G), rok (2008-2013) i sezon wycielenia (wiosna, lato, jesień, zima), liczba urodzonych cieląt (1-2), płeć cielęcia (♀, ♂), liczba zabiegów inseminacyjnych skutkujących ciążą (1-4). Ponadto, w modelowaniu statystycznym przebieg porodu warunkowano cechami związanymi z oceną punktową budowy krów: wysokość w krzyżu, głębokość tułowia, szerokość klatki piersiowej, pozycja zadu, szerokość zadu. Wzięto pod uwagę również długość ciąży i masę cielęcia po urodzeniu, a w przypadku wieloródek także średnią wydajność mleczną za ostatnie 140 dni laktacji (tab. 2).

Tabela 1

Rozkład przebiegu porodu (CE) pierwiastek i wieloródek w zależności od rozpatrywanych czynników

Czynniki	razem	Pierwiastki				razem	Wieloródki			
		CE					CE			
		1	2	3	4		1	2	3	4
Stado										
A	1129	2,92	94,69	2,39	0,00	1734	3,29	95,85	0,81	0,06
B	699	8,58	56,22	25,89	9,30	829	22,20	58,26	16,65	2,90
C	993	35,95	62,24	1,81	0,00	1543	74,66	23,27	2,01	0,06
D	4034	72,88	24,94	2,08	0,10	5745	84,40	14,67	0,85	0,07
E	1079	39,57	52,73	7,69	0,00	1802	53,00	41,29	5,72	0,00
F	1069	1,87	95,88	2,25	0,00	1574	2,60	95,68	1,65	0,06
G	522	0,57	2,68	96,74	0,00	803	1,37	98,38	0,25	0,00
Rok wycielenia										
2008	558	24,55	60,04	15,41	0,00	418	40,43	54,78	4,55	0,24
2009	888	38,63	50,68	10,70	0,00	913	43,59	55,09	1,20	0,11
2010	1455	39,59	48,11	12,30	0,00	1710	47,37	50,99	1,64	0,00
2011	2234	45,17	46,06	8,73	0,04	2953	52,79	46,02	1,12	0,07
2012	2294	44,12	45,20	10,46	0,22	4165	57,38	39,26	3,29	0,07
2013	2096	36,40	54,53	6,06	3,01	3871	49,68	46,22	3,49	0,62
Sezon wycielenia										
wiosna	2562	38,80	49,06	11,44	0,70	3020	47,68	49,50	2,55	0,26
lato	2384	38,55	50,08	10,19	1,17	4064	53,03	43,63	3,13	0,22
jesień	2204	41,06	52,00	6,31	0,64	3746	53,82	43,59	2,30	0,29
zima	2375	43,03	46,19	10,40	0,38	3200	51,19	46,44	2,28	0,09
Liczba cieląt										
1	9473	40,46	49,23	9,57	0,73	13 593	52,67	45,02	2,10	0,20
2	52	13,46	57,69	28,85	0,00	437	20,18	61,24	17,66	0,92
Płeć cielęcia										
♀	5779	45,60	45,75	8,10	0,55	7012	55,22	42,51	2,12	0,14
♂	3740	32,22	54,76	12,03	0,99	6981	48,35	48,65	2,72	0,29
Liczba zabiegów inseminacyjnych										
1	6864	43,75	44,93	10,62	0,70	7815	57,59	40,41	1,80	0,19
2	1665	31,29	59,82	7,99	0,90	3051	53,00	43,56	3,21	0,23
3	585	29,74	62,56	7,35	0,34	1387	37,64	58,33	3,68	0,36
4	411	34,55	60,34	4,14	0,97	1777	34,27	61,40	4,11	0,23

CE: 1 – poród samodzielny, 2 – poród łatwy, 3 – poród trudny, 4 – poród ciężki lub cesarskie cięcie

Przebieg porodu pierwiastek i wieloródek analizowano statystycznie, wykorzystując metodę drzew klasyfikacyjnych. Przystępując do analizy statystycznej dokonano podziału zbioru obserwacji na dwa podzbiory: treningowy (60%) i walidacyjny (40%). Zostały one wygenerowane metodą losowania warstwowego proporcjonalnego. Budując drzewo klasyfikacyjne założono, że minimalna wielkość węzła końcowego nie może być mniejsza niż 30 osobników, zaś głębokość drzewa nie większa niż 6 poziomów. Konstruując drzewo klasyfikacyjne wykorzystano algorytm CART, który w swojej budowie uwzględnia dwa różne kryteria tworzenia podziałów: funkcję entropii (1) i indeks Gini (2) [21].

Funkcja entropii:

$$H(p_1, p_2, ..., p_k) = -\sum_{j=1}^{k} p_j \log_2(p_j) \qquad (1)$$

Indeks Gini (G(p)):

$$G(p) = 1 - \sum_{j=1}^{k} p_j^2 \qquad (2)$$

gdzie:

p – wektor prawdopodobieństwa przynależności obiektów do klas w postaci:

$$p = (p_1, p_2, ..., p_k) = \left(\frac{l_1}{n}, \frac{l_2}{n}, ..., \frac{l_k}{n}\right),$$

k – liczba klas,

l – liczebność klasy,

n – liczebność badanej grupy.

Jakość dopasowania skonstruowanych modeli do danych porównywano za pomocą: średniego błędu kwadratowego,

Tabela 2

Charakterystyka statystyczna przebiegu porodu (CE) pierwiastek i wieloródek

Czynnik	Razem		CE			
			1	2	3	4
Pierwiastki						
wysokość w krzyżu	3822	średnia	83,94	83,53	83,16	83,27
		SD	3,21	3,51	4,02	4,50
głębokość tułowia	3822	średnia	81,27	81,20	80,64	82,45
		SD	2,73	3,01	3,78	3,50
szerokość klatki piersiowej	3822	średnia	79,13	80,38	79,78	81,27
		SD	3,76	3,34	3,61	4,00
pozycja zadu	3822	średnia	79,46	79,74	79,07	79,00
		SD	4,03	3,85	4,10	3,90
szerokość zadu	1917	średnia	85,30	84,68	84,36	85,91
		SD	2,77	2,95	2,85	2,47
długość ciąży	8463	średnia	278,58	278,32	278,02	277,44
		SD	6,05	5,67	5,76	4,54
masa cielęcia	9062	średnia	34,18	34,88	32,90	37,36
		SD	3,70	4,33	2,88	2,62
Wieloródki						
wysokość w krzyżu	4908	średnia	83,85	84,06	83,50	85,50
		SD	2,83	3,50	3,46	2,35
głębokość tułowia	4908	średnia	80,86	80,65	80,24	81,67
		SD	2,32	2,85	2,97	1,97
szerokość klatki piersiowej	4908	średnia	80,54	80,36	80,06	81,83
		SD	2,34	2,55	2,64	2,79
pozycja zadu	4908	średnia	80,06	79,57	80,13	80,17
		SD	3,46	3,32	2,52	2,93
szerokość zadu	320	średnia	86,09	84,72	86,80	86,00
		SD	2,43	3,29	2,49	2,83
długość ciąży	12 876	średnia	279,38	278,94	278,14	280,39
		SD	5,50	6,11	6,53	4,58
masa cielęcia	13 695	średnia	34,98	35,11	35,08	38,44
		SD	4,20	4,66	5,16	4,46
średnia wydajność mleczna za ostatnie 140 dni laktacji	13 215	średnia	23,52	22,46	23,70	23,96
		SD	7,16	7,77	6,74	4,68

CE: 1 – poród samodzielny, 2 – poród łatwy, 3 – poród trudny, 4 – poród ciężki lub cesarskie cięcie
SD – odchylenie standardowe

W tabeli 3. przedstawiono statystyki świadczące o jakości dopasowania modeli służących wskazaniu zmiennych warunkujących przebieg porodu pierwiastek i wieloródek. Do dalszej analizy statystycznej zaproponowano model drzewa zbudowany na podstawie kryterium podziału w postaci indeksu Gini, ponieważ charakteryzował się on korzystniejszymi miarami jakości niż przy zastosowaniu funkcji entropii.

W tabeli 4. podano informację na temat liczby podziałów zbioru danych zrealizowanych na podstawie wyszczególnionych zmiennych oraz ich względne znaczenie (miara Importance) w budowie drzewa decyzyjnego. Otrzymany ranking zmiennych dla pierwiastek i wieloródek wykazał, że najważniejsze znaczenie w tworzeniu drzewa miały: stado, rok wycielenia, płeć i masa ciała cieląt. Przebieg porodu pierwiastek różnicował dodatkowo sezon wycielenia, wysokość w krzyżu i szerokość klatki piersiowej. W przypadku wieloródek zmiennymi różnicującymi były: liczba urodzonych cieląt, długość ciąży, średnia wydajność mleka za ostatnie 140 dni poprzedniej laktacji i głębokość tułowia.

Użyte w budowie drzewa decyzyjnego zmienne (tab. 4) były w znacznej mierze wskazywane przez innych autorów jako źródło zmienności przebiegu porodu. We wcześniejszych badaniach Piwczyńskiego i wsp. [20] zmiennymi, które wpływały na łatwość wycieleń krów były (w kolejności malejącej): kolejna laktacja, masa ciała cielęcia, system utrzymania, okres międzyciążowy, płeć cielęcia oraz długość ciąży. W badaniach Bayram'a i wsp. [2] masa ciała urodzonych cieląt, typ urodzenia i sezon wycielenia miały największy wpływ na przebieg porodu. Autorzy ci wykazali, że trudności z porodem w przypadku porodów bliźniaczych były około cztery razy większe (18,9%) niż w porodach pojedynczych (5,0%). Z kolei w badaniach Zaborskiego i wsp. [23] naj-

Tabela 3

Statystyki dopasowania modelu

Rodzaj statystyki	Indeks Gini	Funkcja entropii
Pierwiastki		
średni błąd kwadratowy	0,0767	0,1221
skumulowany lift	9,8957	9,9843
statystyka Kołmogorowa-Smirnova	0,9360	0,9240
przeciętna funkcja błędu	0,2198	0,4518
pole pod krzywą ROC	0,9890	0,9900
Wieloródki		
średni błąd kwadratowy	0,0594	0,0598
skumulowany lift	9,9941	9,0581
statystyka Kołmogorowa-Smirnova	0,8140	0,9080
przeciętna funkcja błędu	0,1579	0,1634
pole pod krzywą ROC	0,9850	0,9710

przeciętnej funkcji błędu, skumulowanego lift'u, statystyki Kołmogorowa-Simirnova oraz pola pod krzywą ROC. Dwie pierwsze z wyżej wymienionych miar wraz z poprawą jakości modelu dążą do zera, pozostałe zaś wykazują tendencję odwrotną. Analizę statystyczną przeprowadzono za pomocą oprogramowania Enterprise Miner 12.1 należącego do pakietu SAS [21].

Tabela 4

Liczba reguł podziału i miara znaczenia cechy w budowie drzewa

Wyszczególnienie	Liczba reguł podziałów	Miara Importance
Pierwiastki		
stado	5	1
rok wycielenia	7	0,3359
masa cielęcia	5	0,2406
płeć cielęcia	2	0,1523
sezon wycielenia	3	0,1297
wysokość w krzyżu	2	0,0737
szerokość klatki piersiowej	2	0,0339
Wieloródki		
stado	8	1
masa cielęcia	8	0,3039
rok wycielenia	7	0,2607
liczba cieląt	1	0,1409
płeć cielęcia	2	0,1064
długość ciąży	2	0,0392
średnia wydajność mleczna za ostatnie 140 dni laktacji	1	0,0166
głębokość tułowia	1	0,0097

ważniejszymi czynnikami były: długość ciąży, kondycja oraz wiek krowy.

Na rysunkach zaprezentowano diagramy drzew decyzyjnych modelujące przebieg porodu pierwiastek (rys. 1a i 1b) i wieloródek (rys. 2a i 2b), zawierające dane zbioru treningowego. W każdym węźle stanowiącym element składowy drzewa decyzyjnego zawarto informację dotyczącą numeru węzła, procentowego rozkładu przebiegu porodu oraz liczby ocenianych porodów (rys. 1a). Otrzymane projekty drzew decyzyjnych zawierały dla grupy pierwiastek 26 podziałów, kończących się 27 liśćmi, a dla wieloródek 30 podziałów drzewa, zakończonych 31 liśćmi.

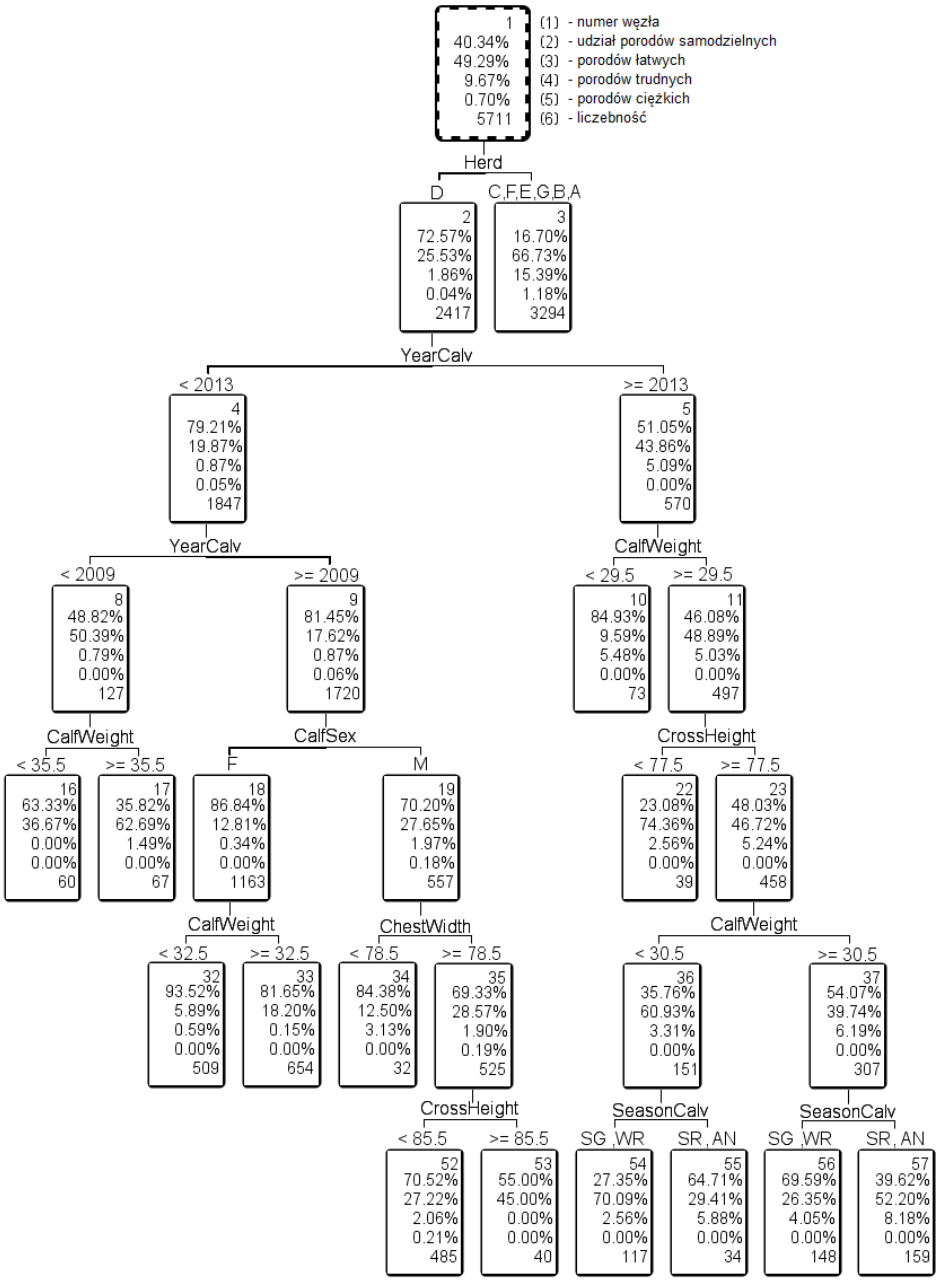
Modelowanie przebiegu porodu pierwiastek (rys. 1a, 1b)

Wielkość zbioru treningowego grupy pierwiastek, tj. liczba analizowanych porodów, wynosiła 5711 (rys. 1a, węzeł 1). Wśród nich 40,34% stanowiły porody samodzielne, 49,29% – łatwe, 9,67% – trudne i 0,70% – ciężkie.

Algorytm odpowiedzialny za budowę drzewa decyzyjnego – CART– wykazał, że najważniejszym czynnikiem różnicującym zbiór pierwiastek w zakresie przebiegu porodu było stado. Na podstawie tej zmiennej dokonał się pierwszy podział zbioru pierwiastek. Do pierwszego podzbioru zostały podporządkowane zwierzęta ze stada D (rys. 1a, węzeł 2), zaś osobniki z pozostałych stad stanowiły podzbiór drugi (węzeł 3). Pierwiastki pochodzące ze stada D charakteryzowały się wyższym o około 56 punktów procentowych (p.p.) udziałem porodów samodzielnych i o 13,5 p.p. niższym ciężkich, w stosunku do rówieśniczek z pozostałych stad (węzeł 3).

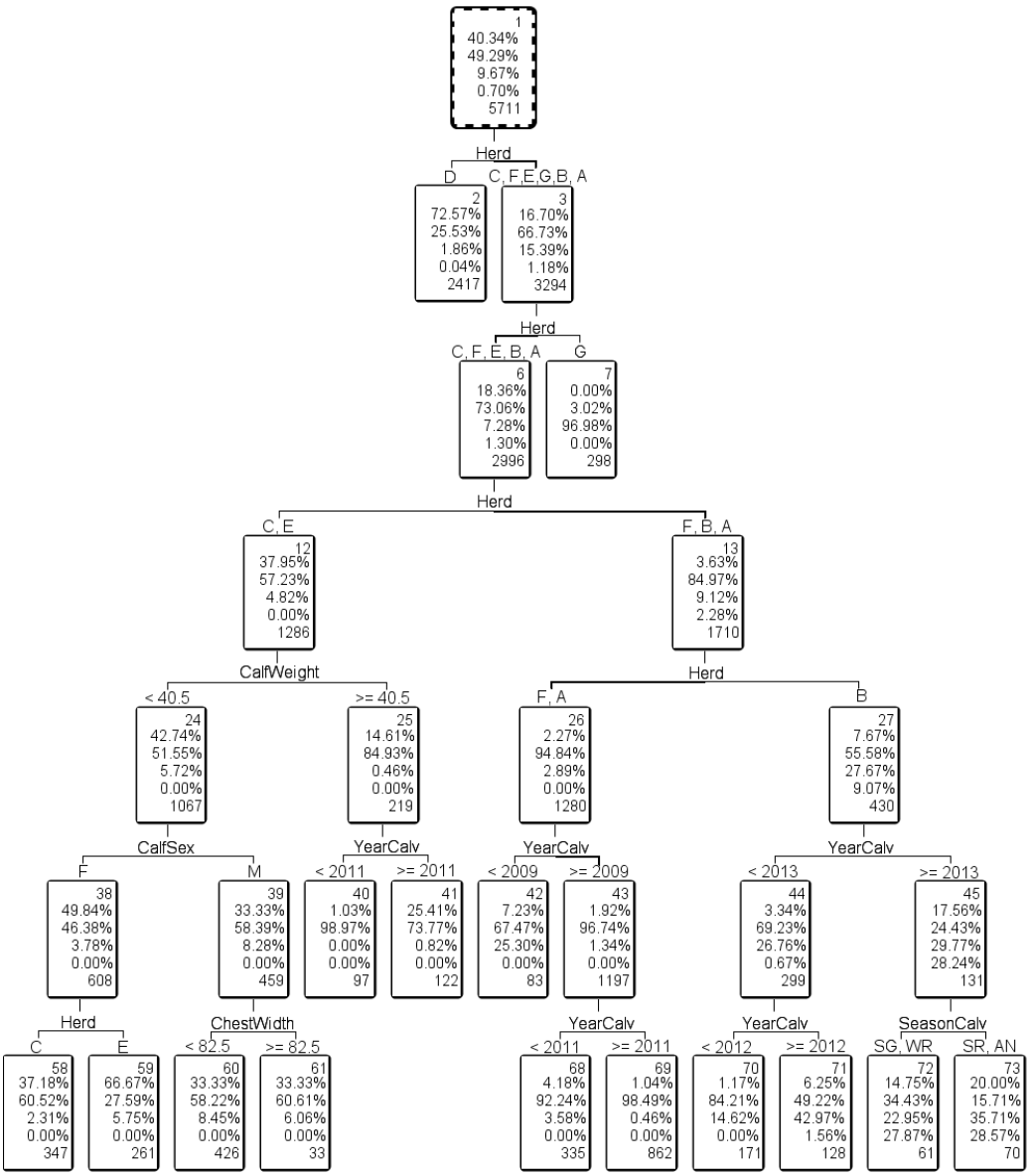
Pierwszą zmienną, dwukrotnie różnicującą przebieg porodów w stadzie D był rok wycielenia. Stwierdzono, że udział porodów samodzielnych był do roku 2012 (węzeł 4) o 28 p.p. większy niż w latach późniejszych (węzeł 5). Następny podział z udziałem zmiennej rok wycielenia grupował zwierzęta ze stada D na urodzone przed (węzeł 8) i po roku 2009 (węzeł 9). Różnica między powyższymi grupami, z korzyścią dla lat wycieleń 2009-2012, wynosiła aż 33 p.p. Wśród zwierząt cielących się przed rokiem 2009 (węzeł 8) jedynym czynnikiem grupującym okazała się masa ciała cieląt po wycieleniu. Wykazano, że w przypadku pierwiastek rodzących cielęta o masie ciała do 35,5 kg 63,3% porodów odbywało się bez asysty człowieka, a 36,7% było porodami łatwymi (węzeł 16). Zaobserwowano również, że przekroczenie progu 35,5 kg masy ciała skutkowało praktycznie odwrotną strukturą porodów (węzeł 17). Z kolei wśród pierwiastek cielących się w latach 2009-2012 zmienną determinującą przebieg porodu była płeć cielęcia. Stwierdzono, że w zwierzęta rodzące cieliczki (węzeł 18) tylko w około 13% wymagały obecności człowieka, podczas gdy rodzące byczki (węzeł 19) w około 30%. Ponadto zaobserwowano, że łatwiejsze porody cieliczek były wtedy, gdy ich masa ciała nie przekroczyła progu 32,5 kg – aż 93,5% porodów samodzielnych. Z kolei w grupie porodów skutkujących urodzeniem byczka ważną zmienną okazała się szerokość klatki piersiowej krów. Stwierdzono, że grupa o szerokości klatki piersiowej poniżej 78,5 pkt. (liść 34) charakteryzowała się o 15 p.p. większym udziałem porodów bez asysty w stosunku do grupy o szerokości klatki piersiowej powyżej tej wartości (węzeł 35).

Przeprowadzona analiza statystyczna wykazała, że czynnikiem, który różnicował łatwość porodu pierwiastek cielących się od roku 2013 w stadzie D (węzeł 5) była masa ciała rodzonego potomstwa. Pierwiastki rodzące cielęta o masie poniżej 29,5 kg wymagały obecności człowieka w około 15% przypadków, podczas gdy porody skutkujące urodzeniem potomstwa o masie ciała równej lub wyższej niż 29,5 kg aż w 54%. Podkreślenia wymaga fakt, że podzbiór złożony z pierwiastek, które rodziły cielęta o masie ciała poniżej 29,5 kg nie podlegał już



Herd – stado, YearCalv – rok wycielenia, CalfWeight – masa cielęcia, CalfSex – płeć cielęcia, F – cieliczki, M – byczki, CrossHeight – wysokość w krzyżu, ChestWidth – szerokość klatki piersiowej, SeasonCalv – sezon wycielenia, SG – wiosna, WR – zima, SR – lato, AN – jesień

Rys. 1a. Drzewo klasyfikacyjne pierwiastek (część 1)



Rys. 1b. Drzewo klasyfikacyjne pierwiastek (część 2)

dalszym podziałom i stał się liściem (liść 10). Zatem dalszym podziałom podlegał zbiór samic rodzących potomstwo o wyższej masie ciała (węzeł 11) – pierwszy podział dokonał się na podstawie wysokości w krzyżu. Rodzące samice o niższej wysokości w krzyżu (liść 22) częściej wymagały asysty człowieka (77%) niż o wysokości wyższej (węzeł 23; 52%) – różnica około 25 p.p. Ten ostatni zbiór został podzielony ze względu na masę ciała potomstwa – poniżej 30,5 kg (węzeł 36) oraz równej lub wyższej (węzeł 37) tej wartości. Obie wyżej wymienione grupy zwierząt zostały podzielone ze względu na porę roku (liście 54, 55, 56 i 57). Jednak efekt poszczególnych pór roku na przebieg porodu był niejednoznaczny – był warunkowany masą ciała rodzonego potomstwa.

W analogiczny sposób należy interpretować zbiór pierwiastek użytkowanych w stadach C, F, E, G, B, A (rys. 1b, węzeł 3). Zaobserwowano, że najważniejszym czynnikiem wpływającym na przebieg porodu było stado – klasyfikując zwierzęta aż czterokrotnie. Świadczyć to może o zróżnicowanych warunkach utrzymania i żywienia zwierząt w poszczególnych stadach. Zmienna rok wycielenia została aż pięciokrotnie wykorzystania w klasyfikowaniu pierwiastek z wyżej wymienionych stad w zakresie przebiegu porodu. Z kolei zmienne: płeć i masa ciała rodzących się cieląt, sezon wycielenia oraz szerokość klatki piersiowej uczestniczyły w grupowaniu pierwiastek jednokrotnie.

Podsumowując graficzny model drzewa prezentujący rozkład przebiegu porodu pierwiastek w stadach A, B, C, E, F i G należy stwierdzić, że najtrudniejsze porody miały miejsce w stadzie B, wśród samic cielących się od 2013 roku w sezonie letnio-jesiennym – aż 28,24% porodów trudnych i ciężkich (węzeł 45). Z kolei najwięcej porodów samodzielnych i łatwych obserwowano wśród zwierząt ze stad C i E, które rodziły potomstwo o masie ciała powyżej 40,5 kg przed 2011 rokiem (węzeł 40).

Modelowanie przebiegu porodu wieloródek (rys. 2a, 2b)

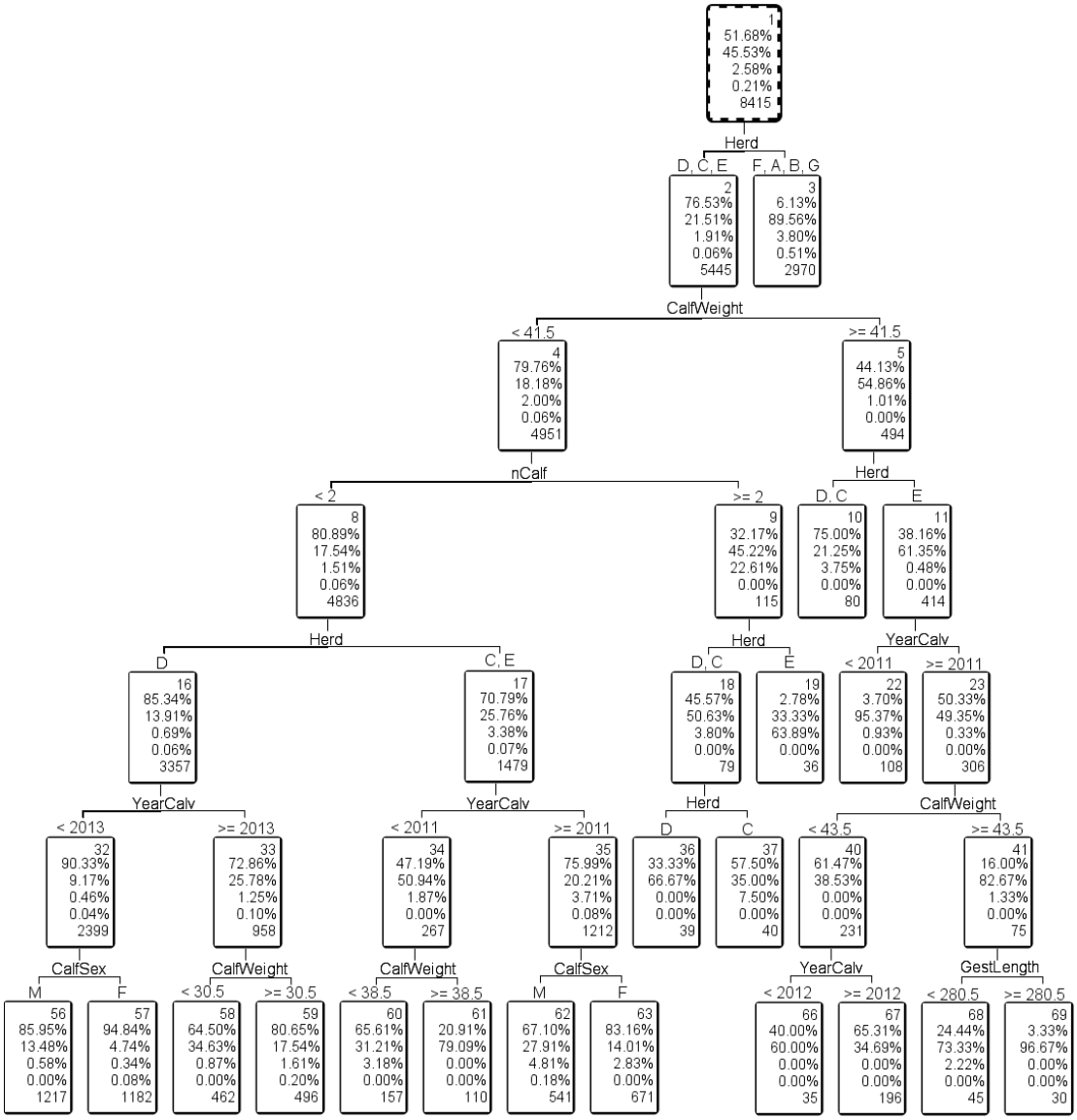
Zmiennymi, które najczęściej różnicowały zbiór wieloródek były: stado i masa cielęcia po urodzeniu (po 8 podziałów) oraz rok wycielenia (7 podziałów). Na uwagę zasługują jednak również pozostałe zmienne uczestniczące w tworzeniu drzewa decyzyjnego. Zaobserwowano, że w przypadku ciąż pojedynczych było więcej o około 49 p.p. samodzielnych (węzeł 8), a mniej o 21 p.p. trudnych i ciężkich przebiegów porodu w porównaniu z mnogimi (węzeł 9). Biorąc pod uwagę płeć cieląt stwierdzono, że wycielenia cieliczek były łatwiejsze w porównaniu z buhajkami – różnica w udziale porodów samodzielnych wahała się od 9 (węzły 56 i 57) do 16 p.p. (węzły 62 i 63). Zauważenia wymagają porody z ciąż trwających powyżej 283 dni – aż w 12,9% porodów miało trudny przebieg (węzeł 49). Warto podkreślić, iż wieloródki ze stad F i G (węzeł 42) produkujące dziennie powyżej 21,3 kg mleka cieliły się łatwiej (12,5% porodów samodzielnych i 87,5% łatwych) w porównaniu z krowami dającymi mniejszą ilość mleka – 98% porodów łatwych oraz 1,79% ciężkich.

Wzrost ryzyka porodów z asystą człowieka w grupie pierwiastek w badaniach własnych został wykazany we wcześniejszych badaniach innych autorów [10, 13, 14, 16, 18]. W badaniach Lombarda i wsp. [10] więcej niż połowa porodów pierwiastek (51,2%), w porównaniu z 29,4% porodami wieloródek, wymagała asysty człowieka. Również Olson i wsp. [16] wskazują, że pierwiastki rodziły cielęta o niższej masie ciała, ich ciąża była krótsza i wymagały około 2,5 razy częściej pomocy przy porodzie. W badaniach Meyer i wsp. [15] grupa pierwiastek charakteryzowała się ponad trzykrotnie wyższym ryzykiem problemów z porodem w stosunku do krów wieloródek.

W badaniach własnych około 49% zarejestrowanych porodów pierwiastek i 45% wieloródek odbywało się z asystą człowieka. W badaniach Piwczyńskiego i wsp. [19] około jedna trzecia porodów odbywała się z asystą człowieka. Zdaniem Mee [11], większa liczba porodów samodzielnych w oborach możliwa

jest do uzyskania jedynie dzięki dobrej opiece oborowych i dokładnej obserwacji wycieleń.

Meyer i wsp. [14] wskazali, że w grupie krów pierwiastek sezon i rok urodzenia oraz długość ciąży były czynnikami najlepiej prognozującymi dystocję. Z kolei Bayram i wsp. [2] za czynniki mające największy wpływ na trudności z porodem uznali masę urodzonych cieląt, typ urodzenia i sezon wycielenia krowy. Również we wcześniejszych badaniach autorów [20] wykazano, że zaproponowany model drzewa klasyfikacyjnego okazał się dogodnym narzędziem do wyodrębnienia czynników odpowiedzialnych za przebieg porodu krów. Otrzymane wyniki pozwoliły stwierdzić, że najbardziej złożony przebieg porodu miał miejsce w grupie pierwiastek, które rodziły cielęta o wysokiej masie ciała. Najłatwiejsze porody wystąpiły w populacji wieloródek rodzących potomstwo o masie ciała do 35,5 kg i wycielonych do 282. dnia ciąży.



Herd – stado, YearCalv – rok wycielenia, CalfWeight – masa cielęcia, CalfSex – płeć cielęcia, F – cieliczki, M – byczki, CrossHeight – wysokość w krzyżu, ChestWidth – szerokość klatki piersiowej, SeasonCalv – sezon wycielenia, SG – wiosna, WR – zima, SR – lato, AN – jesień, GastLenght – długość ciąży, nCalf – liczba cieląt, avgMilkCalv140 – wydajność mleczna za ostatnie 140 dni laktacji, TrunkDepth – głębokość tułowia

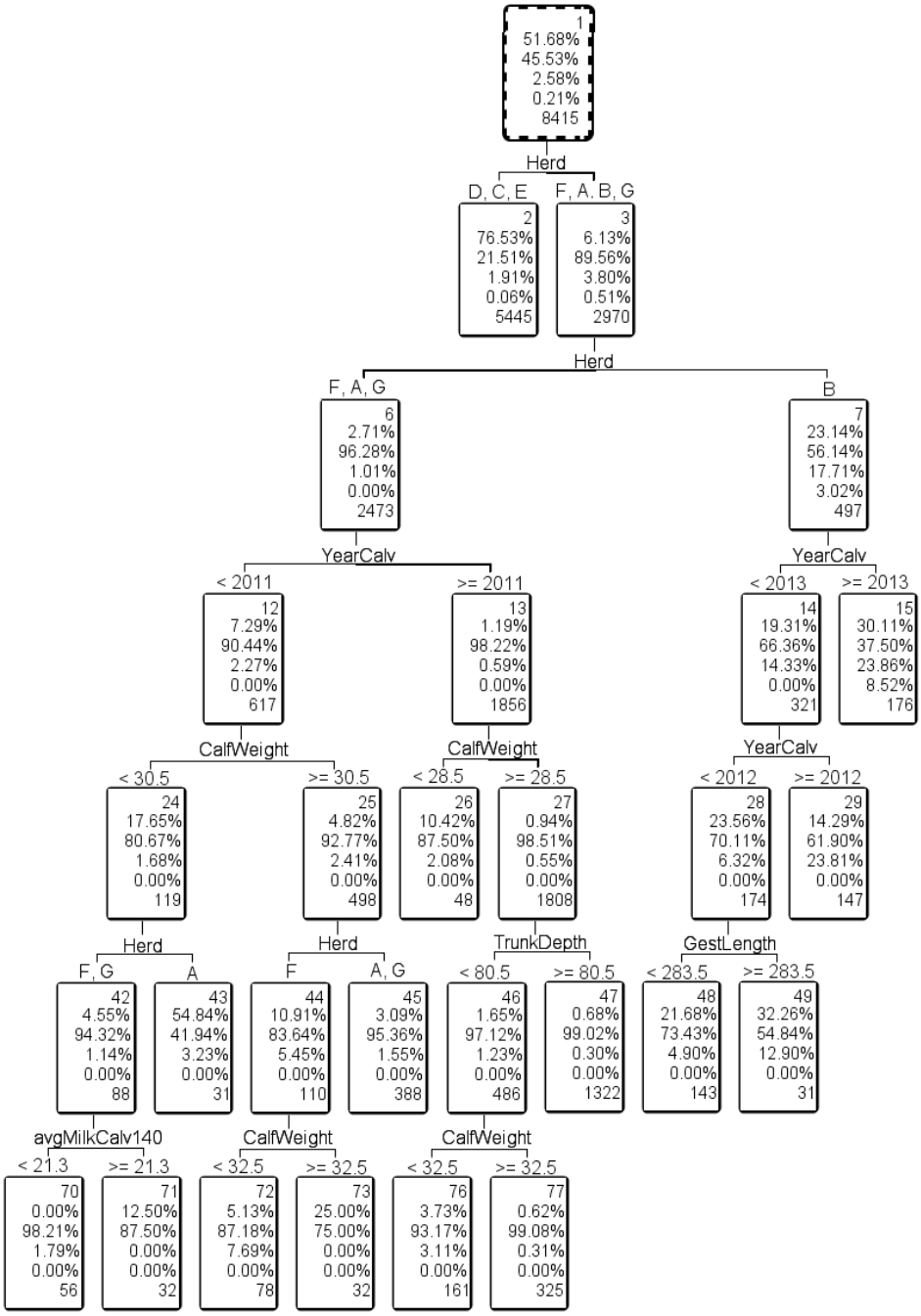
Rys. 2a. Drzewo klasyfikacyjne wieloródek (część 1)

Podsumowanie

Podsumowując wyniki przeprowadzonych badań wykazano, że zarówno w grupie pierwiastek, jak i wieloródek czynnikami różnicującymi zwierzęta w zakresie przebiegu porodu były: stado, rok wycielenia, płeć oraz masa ciała urodzonego cielęcia. Wśród pierwiastek dodatkowe podziały spowodowane były na podstawie zmiennych: sezon wycielenia, wysokość w krzyżu i szerokość klatki piersiowej, zaś w przypadku wieloródek – długość ciąży, liczba urodzonych cieląt, średnia wydajność mleczna za ostatnie 140 dni laktacji i głębokość tułowia. Ponadto wykazano, że mimo skonstruowania drzew klasyfikacyjnych modelujących przebieg porodu trudno jednoznacznie określić znaczenie czynników, ze względu na występowanie licznych interakcji między nimi.

Cechy, które wprowadzono do modelu, związane z oceną budowy krów oraz bezpośrednio z rozrodem nie stanowiły podstawowych kryteriów podziału drzew. Pojawiały się dopiero w ostatnich podziałach, gdzie liczebność krów w poszczególnych grupach była już mniejsza.

Literatura: 1. Banos G., 1996 – Survey on genetic evaluation procedures for functional traits in cattle in various countries. Proceedings of the International Workshop on Genetic Improvement of Functional Traits in Cattle. Interbull Bulletin 12, 11-24. **2. Bayram B., Topal M., Aksakal V., Önk K.**, 2015 – Investigate the Effects of Non-genetic Factors on Calving Difficulty and Stillbirth Rate in Holstein Friesian Cattle Using the CHAID Analysis. Kafkas Üniversitesi Veteriner Fakültesi Dergisi 21, 5, 645-652. **3. Berry D.P., Lee J.M., Macdonald K.A., Roche J.R.**, 2007 – Body Condition Score and Body Weight Effects on Dystocia and Stillbirths and Consequent Effects on Postcalving Performance. J. Dairy Sci. 90, 4201-4211. **4. Bicalho R.C., Galvao K.N., Warnick L.D., Guard C.L.**, 2008 – Stillbirth parturition reduces milk production in Holstein cows. Prev. Vet. Med. 4, 84, 112-120. **5. De Vries A.**, 2006 – Economic Value of Pregnancy in Dairy Cattle. J. Dairy Sci. 89, 3876-3885. **6. Grzesiak W., Zaborski D., Sablik P., Pilarczyk R.**, 2011 – Detection of difficult conceptions cows using selected data mining methods. Anim. Sci. Pap. Rep. 29, 4, 293-302. **7. Gundelach Y., Essmeyer K., Teltscher M.K., Hoedemaker M.**, 2009 – Risk factors for perinatal mortality in dairy cattle: Cow and foetal factors, calving process. Theriogenol. 71, 901-909. **8. Hansen M., Lund M.S., Pedersen J., Christensen L.G.**, 2004 – Gestation length in Danish Holsteins has weak genetic associations with stillbirth, calving difficulty, and calf size. Liv. Prod. Sci. 91, 23-33. **9. Krzyżewski J., Reklewski Z.**, 1997 – Chów i hodowla zwierząt gospodarskich. Fundacja Rozwój SGGW. **10. Lombard J.E., Garry F.B., Tomlinson S.M., Garber L.P.**, 2007 – Impacts of dystocia on Health and Survival of Dairy Calves. J. Dairy Sci. 90, 4, 1751-1760. **11. Mee J.F.**,



Rys. 2b. Drzewo klasyfikacyjne wieloródek (część 2)

2004 – Managing the dairy cow at calving. Veterinary Clinics of North America, Food Animal Practice 20, 521-546. **12. Mee J.F., Berry D.P., Cromie A.R.**, 2008 – Prevalence of, and risk factors associated with, perinatal calf mortality in pasture-based Holstein-Friesian cows. Animal 2, 4, 613-620. **13. Mee J.F., Berry D.P., Cromie A.R.**, 2011 – Risk factors for calving assistance and dystocia in pasture-based Holstein-Friesian heifers and cows in Ireland. Vet. J. 187, 189-194. **14. Meyer C.L., Beger P.J., Koehler K.J.**, 2000 – Interactions among factors affecting stillbirths in Holstein cattle in the United States. J. Dairy Sci. 83, 2657-2663. **15. Meyer C.L., Berger P.J., Koehler K.J., Thompson J.R., Sattler C.G.**, 2001 – Phenotypic trends in incidence of stillbirth for Holsteins in the United States. J. Dairy Sci. 84, 515-523. **16. Olson K.M., Cassell B.G., McAllister A.J., Washburn S.P.**, 2009 – Dystocia, stillbirth, gestation length, and birth weight in Holstein, Jersey, and reciprocal crosses from a planned experiment. J. Dairy Sci. 92, 6169-6175. **17.** PFHBiPM, 2015 – Ocena i Hodowla Bydła Mlecznego. **18. Philipson J.**, 1996 – Strategies to reduce problems in calving performance and stillbirths by selection and differential use of bulls. Interbull Bulletin 12, 65-71. **19. Piwczyński D., Sitkowska B., Wiśniewska E.**, 2012 – Application of classification trees and logistic regression to determine factors responsible for lamb mortality. Small Ruminant Res. 103, 2, 225-231. **20. Piwczyński D., Nogalski Z., Sitkowska B.**, 2013 – Statistical modeling of calving ease and stillbirths in dairy cattle using the classification tree technique. Livestock Sci. 154, 19-27. **21.** SAS Institute Inc., 2014 – Getting started with SAS® Enterprise Miner 7.1TM. Cary, NC: SAS Institute Inc. **22. Uematsu M., Sasaki Y., Kitahara G., Semeshima H., Osawa T.**, 2013 – Risk factors for stillbirth and dystocia in Japanese Black cattle. Vet. J. 198, 212-216. **23. Zaborski D., Grzesiak W., Kotarska K., Szatkowska I., Jedrzejczak M.**, 2014 – Detection of difficult calvings in dairy cows using boosted classification trees. Indian J. Anim. Res. 48, 5, 452-458. **24. Zaborski D., Grzesiak W., Pilarczyk R.**, 2014 – Detection of difficult calvings in the Polish Holstein-Friesian Black-and-White heifers. J. Appl. Anim. Res. 44, 1, 42-53.

Modelling of the course of parturition using classification trees

Summary

The objective of this study was to statistically model calving ease in primiparous and multiparous Holstein-Friesian cows using classification trees. Calving ease was coded on a four-point scale: 1 – unassisted calving; 2 – easy calving; 3 – difficult calving; 4 – very difficult calving or Caesarean section. Statistical models were built using the CART algorithm and two criteria for division of the data set, i.e. the entropy function and the Gini index. The statistical analysis revealed that in both the primiparous and multiparous groups the factors which most frequently differentiated nodes (five to eight times) in both trees were the herd, the year of calving, the body weight of the calf, and its sex. Despite the construction of classification trees to model calving ease, it is difficult to clearly determine the effect of factors related to reproduction and conformation of cows, as the model contains too many other important criteria.

KEY WORDS: dairy cattle, calving ease, classification trees