

Estimativa de estro em vacas leiteiras utilizando métodos quantitativos preditivos

Dairy cows estrus estimation using predictive and quantitative methods

Irenilza de Alencar Nääs^{I*} Marcos Paulo Garcia Queiroz^I Daniella Jorge de Moura^I
Leandro dos Anjos Brunassi^{II}

-NOTA-

RESUMO

O Brasil é o sexto maior produtor de leite do mundo, sendo que essa produção cresce a uma taxa anual 4% superior aos demais países produtores. Parte desse aumento na produção de leite deve-se ao uso de diversas tecnologias desenvolvidas para o setor, principalmente, aquelas relacionadas à genética e ao manejo do rebanho. A detecção acurada do cio em vacas é um fator limitante na eficiência reprodutiva do rebanho leiteiro, sendo considerada uma das principais deficiências na reprodução bovina. Falha na identificação do estro com eficiência ocasiona perdas para o produtor. Métodos quantitativos preditivos, baseados em dados históricos e conhecimento especialista, permitem, a partir de uma base de dados organizada, a predição de padrões com baixa percentagem de erro. Este trabalho comparou a precisão das técnicas de estimativa de estro para vacas da raça Holandesa alojadas em galpão freestall, utilizando métodos quantitativos preditivos, por meio da interposição dos pontos intermediários provenientes de série histórica do rebanho. Uma base de regras foi formulada sendo que os valores dos pesos de cada afirmação pertencem a um intervalo de zero a um, e esses limites foram utilizados para gerar a função de pertinência Fuzzy, cuja saída era a predição de estro. Na etapa seguinte, foi aplicada a técnica de Data mining utilizando os parâmetros de movimentação, produção de leite, dias de lactação e comportamento de monta, sendo gerada uma árvore de decisão para analisar os parâmetros mais significativos na previsão de estro em vacas leiteiras. Os resultados indicaram que a presença de estro pode ser detectada com maior precisão usando a observação de movimentação das vacas (87%, erro estimado 4%) ou o comportamento de monta (78%, erro estimado 11%).

Palavras-chave: modelagem preditiva, mineração de dados, lógica Fuzzy.

ABSTRACT

Brazil is the sixth world's larger milk producer, increasing its production at an annual rate of 4% above other producer countries. Part of this raise in milk production was due to the use of several technologies that have being developed for the sector, mainly those related to genetics and herd management. Accurate estrus detection in dairy cows is a limiting factor in the reproduction efficiency of dairy cattle, and it has been considered the most important deficiency in the field of reproduction. Failing to detect estrus efficiently may cause losses for the producer. Quantitative predictive methods based on historical data and specialist knowledge may allow, from an organized data base, the prediction of estrus pattern with lower error. This research compared the precision of the estrus prediction techniques for freestall confined Holstein dairy cows using quantitative predictive methods, through the interpolation of intermediate points of historical herd data set. A base of rules was formulated and the values of weight for each statement is within the interval of 0 to 1; and these limits were used to generate a function of pertinence fuzzy that had as output the estrus prediction. In the following stage Data mining technique was applied using the parameters of movement rate, milk production, days of lactation and mounting behavior, and a decision tree was built for analyzing the most significant parameters for predicting estrus in dairy cows. The results indicate that the prediction of estrus incidence may be achieved either using the association of cow's movement (87%, with estimated error of 4%) or the observation of mounting behavior (78%, with estimated error of 11%).

Key words: predictive modeling, data mining, Fuzzy logic.

^IFaculdade de Engenharia Agrícola, Universidade Estadual de Campinas (UNICAMP), 13083-970, Campinas, SP, Brasil. E-mail: irenilza@agr.unicamp.br. *Autor para correspondência.

^{II}Programa de Pós-graduação em Engenharia Agrícola, Faculdade de Engenharia Agrícola, UNICAMP, Campinas, SP, Brasil.

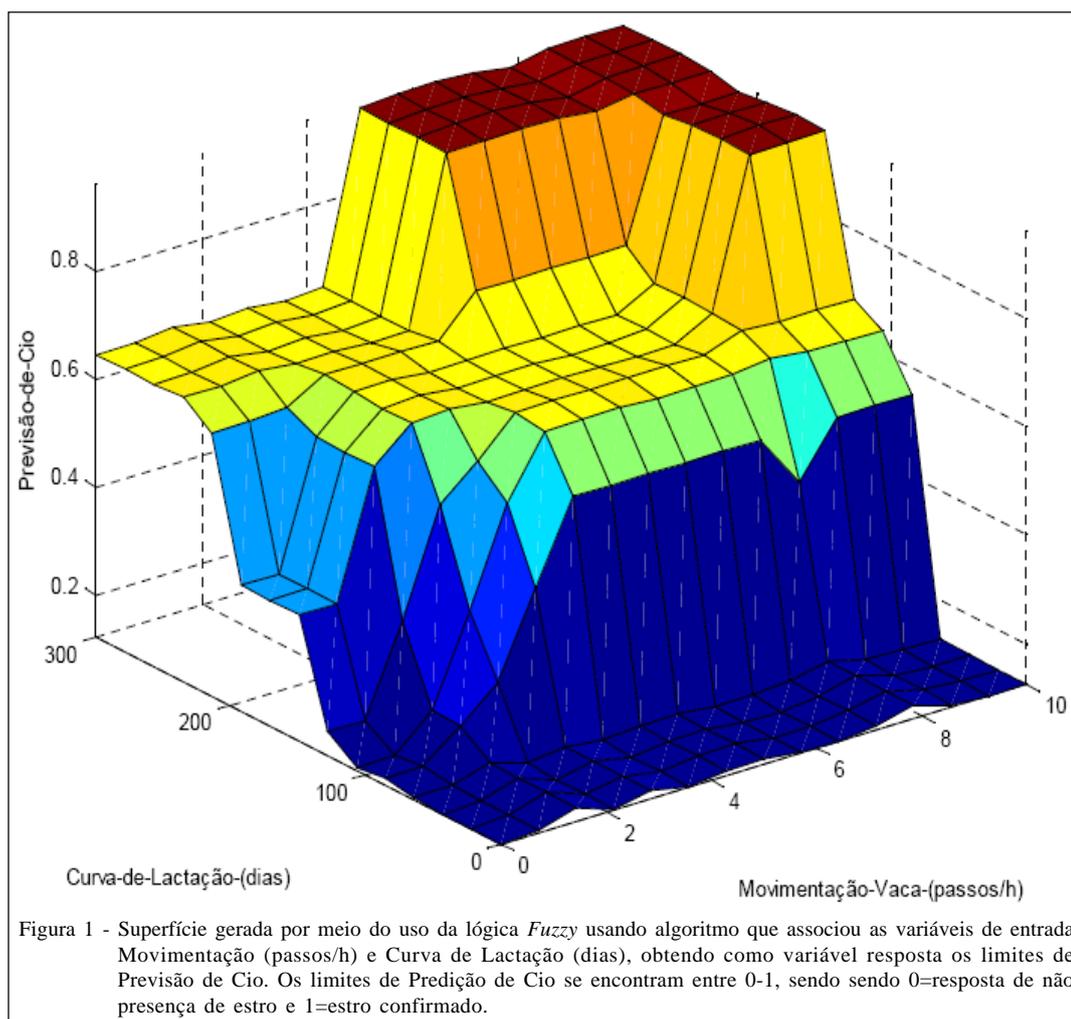
Das técnicas desenvolvidas para detecção de estro em vacas, a principal é a observação visual, que está diretamente relacionada ao compromisso dos responsáveis por checar esta condição, por meio do entendimento dos comportamentos e dos sinais (DISKIN & SCREENAM, 2000). ÁVILA PIRES et al. (2003) constataram que a observação contínua do rebanho eliminou a possibilidade de manifestações de estro não-identificadas, entretanto, segundo HANSEN (2003), é difícil, mesmo para um tratador experiente, identificar mais de 80% das vacas em cio de um rebanho de médio porte, mesmo observando as fêmeas três vezes por dia, por, pelo menos, 30min. Esta técnica, utilizada isoladamente, é geralmente ineficiente, recomendando-se a associação com tecnologias mais eficazes para detecção de estro (KASTELIC, 2001; FIRK et al., 2002; RORIE et al., 2002). As técnicas eletrônicas disponíveis baseiam-se em medidas de movimentação, mudanças na resistência elétrica nas secreções do trato reprodutivo e sensores de monta (RORIE et al., 2002).

Para a estimativa de situações com alto grau de aleatoriedade, como o estro bovino, baseada em análise empírica de medidas relacionadas a eventos e/ou fisiológicas, pesquisas apontam o potencial do uso de critérios específicos em uma base de conhecimento prévio, utilizando a Teoria dos Conjuntos *Fuzzy* (ERADUS et al., 1998; FIRK et al., 2002), que se caracteriza pela incerteza existente entre as afirmações de “sim” e “não” (BEZDEK & SANKAR, 1992). A mineração de dados é uma etapa do procedimento descrito como *Knowledge Discovery in Databases* (KDD), que é um processo de identificação nos dados de padrões válidos, potencialmente úteis e compreensíveis (FAYYAD & STOLORZ, 1996). Na fase de mineração (*Data mining*), algoritmos de extração de padrões são aplicados sobre os dados, para serem eliminadas as inconsistências (BERSON & SMITH, 1997).

Neste trabalho foi avaliada a eficácia (em termos de precisão e erro) das técnicas utilizadas para a predição de estro, assim como a acuracidade do estro identificado, utilizando funções de pertinência *Fuzzy*, associadas ao processo de mineração de dados (KDD), para vacas leiteiras da raça Holandesa alojadas em sistemas de *freestall*. Foram escolhidas aleatoriamente no rebanho 100 vacas com características semelhantes e estando na mesma faixa da curva de lactação. Foi aplicada a ANOVA, para avaliar a hipótese de semelhança entre as vacas em produção, que se

confirmou com o teste de comparações múltiplas (teste Tukey, $P \leq 0,05$). Todas as vacas usavam um pedômetro (equipamento eletrônico que registra continuamente a movimentação de animais pela contagem de passos) amarrado à pata dianteira, para sua identificação e avaliação de sua movimentação.

A predição de estro foi efetuada utilizando duas análises: a primeira foi a aplicação da Lógica *Fuzzy*, e a segunda utilizou a técnica de mineração de dados (*Data mining*). Para o desenvolvimento da função de pertinência *Fuzzy*, foi elaborada uma base de regras, com registros retirados do banco de dados de informações históricas do rebanho, que foi organizada em uma planilha que continha as medidas de movimentação, produção de leite, dias em lactação e presença do comportamento de monta, para analisar os atributos mais significativos na determinação do estro de vacas leiteiras. Nessa planilha os registros de movimento (passos h^{-1}) foram divididos em faixas de (<1,5; 1,5–3,0; 3,0–4,5; 4,5–8,0; >8,0), com as seguintes denominações: *muito baixa* (MB), *baixa* (B), *média* (M), *alta* (A), e *muito alta* (MA), respectivamente, conforme informações da literatura (KIDDY, 1977; AT-TARAS & SPAHR, 2001; RORIE et al., 2002; HANSEN, 2003). Os registros de produção de leite ($kg\ dia^{-1}$) foram divididos em faixas de <25, 25–32, e >32, com as seguintes denominações: *baixa* (B), *média* (M) e *alta* (A), respectivamente. Os parâmetros de dias em lactação, compatíveis com a curva de lactação das vacas, foram divididos em faixas de 1–80, 80–120, 120–180, 180–240, e 240–300, com as seguintes denominações, utilizadas usualmente no manejo do rebanho estudado: *inicial* (In), *pico* (P), *pós-pico* (PP), *intermediário* (I) e *final* (F), respectivamente. A presença do comportamento de monta foi estabelecida em padrão binário: sim (S) ou não (N). O método de inferência utilizado foi o Mandani, que combinou o grau de pertinência referente a cada um dos valores de entrada, por meio do operador mínimo, e agregou as regras por meio do operador máximo, tendo sido usadas funções de pertinências do tipo trapezoidal e a saída foi o conjunto *Fuzzy*. A *defuzzificação* foi realizada pelo método do centro de gravidade, permitindo a possibilidade de simular distintas situações entre as variáveis de entrada e ter uma resposta otimizada da variável de saída. A figura 1 mostra a superfície gerada usando o algoritmo *Fuzzy*, que associou as variáveis de entrada Movimentação (passos h^{-1}) e Curva de Lactação (dias), obtendo como variável resposta os limites de Previsão de Cio.



Usando a técnica de *Data mining*, foram estudados os parâmetros de movimentação (passos h⁻¹), produção de leite (kg dia⁻¹), dias de lactação (dia) e presença de monta (S, N) e foi elaborada uma tabela com os dados do subconjunto de 15 vacas do conjunto total das 100 vacas do experimento, para análise na previsão de cio. Os dados obtidos foram submetidos a uma mineração, sendo fundamentada na construção de uma árvore de decisão, de acordo com BACKER (1995), para analisar os parâmetros mais significativos na previsão de estro em vacas leiteiras. Na mineração de dados, foram associados os parâmetros por meio de arquivos tipo *.data* para agregar os dados envolvidos, enquanto um arquivo *.names* foi gerado para descrição dos atributos. Foram inseridos no sistema os dados históricos de 15 vacas, subconjunto do conjunto das 100 vacas, como segue: (Movimentação, passos h⁻¹; Produção de Leite, kg dia⁻¹; Curva de Lactação, dia; e

Comportamento de Monta, sim ou não). A validação do algoritmo foi feita utilizando o conjunto das 100 vacas. As Equações 1 e 2 foram utilizadas para calcular a medida de acuracidade das respostas estimadas de estro, a partir da árvore de decisão adotada, usando o conceito de classificação proposto por EARDLEY et al. (2002).

$$\text{Erro (\%)} = [1 - (\text{número de IC}) / (\text{número de VPC})] \times 100$$

Eq. 1

em que: IC = inferências de cio e VPC = vacas com prenhez confirmada.

$$\text{Precisão (\%)} = (\text{número de IC}) / (\text{número de VEC}) \times 100$$

Eq. 2

em que: VEC = vacas estimadas como no cio.

Na validação, utilizando funções de pertinência Fuzzy, foram encontrados os valores de IC=39, VPC=55 e VEC=63, resultando em erro de 29% e precisão de 62% na estimativa de estro. Embora dentro

de uma faixa de estimativa ampla, esses resultados concordam com o relato de FIRK et al. (2002), que apontam resultados mais precisos na detecção de estro em vacas, com a combinação variáveis possíveis de serem registradas. Na mineração de dados, cada animal foi representado com um vetor, qualificado com um valor binário de: sim ou não, com relação à presença de estro. Cada vetor tinha um atributo em pares para o treino e o algoritmo calculava a informação gerada, de maneira a cada par contribuir para a escolha do maior valor a ser determinado, como o primeiro segmento da árvore de decisão (FAYYAD & STOLORZ, 1996; EARDLEY et al., 2002). Com a utilização da mineração de dados dos parâmetros envolvidos no estro e vacas leiteiras, foi gerado o primeiro resultado, em que foi possível prever que, na árvore de decisão, o parâmetro detecção do cio mais significativo foi o comportamento de monta (IC=49, VPC= 55 e VEC= 63). Esse resultado concorda com a literatura como o parâmetro mais importante na indicação de estro (DISKIN & SCREENAM, 2000; HANSEN, 2003; ÁVILA PIRES et al., 2003). Entretanto, quando se calcula o erro na estimativa após a validação do algoritmo, este se apresenta relativamente alto, em que IC= (11%) e a precisão relativamente é baixa (78%).

Das 100 vacas analisadas na fazenda, 51 vacas aceitaram a monta e estavam no cio, enquanto as outras 49 não aceitaram a monta, porém, destas, 45 não estavam no cio e quatro estavam no cio. Desse modo o fator monta, se usado isoladamente, pode ser limitante para tomada de decisão na previsão de estro. Dados semelhantes foram encontrados por NEBEL et al. (2000) e STURMAN et al. (2000). Na etapa seguinte de processamento, foi excluído o parâmetro presença de monta e foi realizado um novo cálculo para prever o segundo parâmetro mais significativo na detecção de cio (ADRIAANS & ZANTINGE, 1996). Foi possível verificar que o segundo parâmetro mais importante na árvore de decisão foi movimentação (IC=53,

VPC=55 e VEC=63), concordando com o que se encontra descrito na literatura (KIDDY, 1977; KASTELIC, 2001; RORIE et al., 2002). Este resultado mostrou menor erro na estimativa (4%) e maior precisão (87%).

Nos resultados do processamento, foi encontrado que 12 vacas com movimentação muito baixa, não estavam no cio, 25 vacas com movimentação baixa também não estavam no cio e, em oito vacas com movimentação média, seis não estavam no cio, enquanto duas estavam no cio. Por outro lado, em 16 vacas com movimentação alta, somente 14 estavam no cio e, finalmente, das 39 vacas restantes com movimentação muito alta, todas estavam no cio. Foi confirmada que a movimentação das vacas esteve diretamente relacionada ao estro, concordando com HANSEN (2003). Um terceiro parâmetro importante na árvore de decisão foi relacionado à curva de lactação (dia), porém, sua validação apresentou menor precisão (<30%) na detecção do cio nas vacas leiteiras do experimento.

Na tabela 1 encontram-se as comparações com os métodos de estimativa de estro, considerando as Equações 1 e 2, em que é possível observar que essas técnicas quantitativas preditivas permitiram associar todos os dados disponíveis na produção, de maneira a tornar o processo de determinação do estro mais preciso, fundamentando melhor a tomada de decisão do momento ideal para inseminação artificial e reduzindo as perdas já enunciadas por HEERSCHER JR & NEBEL (1994), NEBEL et al. (2000) e FRENCH & NEBEL (2003), aumentando, conseqüentemente, a eficiência da produção leiteira.

AGRADECIMENTO(S)

Os autores agradecem à Embrapa, pelo financiamento do projeto PRODETAB n. 056-02; à Fazenda Campestre, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelas bolsas de PIBIC-IC dos alunos envolvidos no projeto.

Tabela 1 - Comparação entre os algoritmos resultantes da aplicação de métodos quantitativos preditivos em técnicas de estimativa de estro para vacas leiteiras.

| Sistema | Erro (%) | Precisão (%) |
|--|------------|--------------|
| Lógica Fuzzy | (39/55) 29 | (39/63) 62 |
| Data mining utilizando os dados de movimentação | (53/55) 4 | (55/63) 87 |
| Data mining utilizando os dados de curva de lactação | >20 | <30 |
| Data mining utilizando o comportamento de monta | (49/55) 11 | (49/63) 78 |

REFERÊNCIAS

- ADRIAANS, P.; ZANTINGE, D. **Data mining**. Harlow: Addison-Wesley, 1996. 145p.
- AT-TARAS E.E.; SPAHR, S.L. Detection and characterization of estrus in dairy cattle with an electronic heatmount detector and an electronic activity tag. **Journal of Dairy Science**, v.84, n.4, p.792-798, 2001.
- ÁVILA PIRES, M.F. et al. Comportamento de vacas da raça Gir (*Bos taurus indicus*) em estro. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.55, n.2, p.187-196, 2003.
- BACKER, E. **Computer-assisted reasoning in cluster analysis**. New York: Prentice-Hall, 1995. 400p.
- BERSON, A.; SMITH, S.J. **Data warehousing, data mining and OLAP**. New York: McGraw-Hill, 1997. 640p.
- BEZDEK, J.C.; SANKAR, K.P. **Fuzzy models for pattern recognition**. Nova Iorque: IEEE, 1992. 245p.
- DISKIN, M.G.; SCREENAM, J.M. Expression and detection of estrus in cattle. **Reproduction Nutrition Development**, v.40, n.5, p.481-491, 2000.
- EARDLEY, D.D. et al. Decision tree classifiers to confirm pneumonia diagnosis. In: THE AMERICAN MEDICAL INFORMATICS ASSOCIATION SYMPOSIUM, 2002, Scottsdale, Estados Unidos. **Anais...** Scottsdale, Unites States of America: American Medical Informatics Association, 2002. p.1-5. Capturado em: 12 dez. 2004. Online. Disponível na Internet: <http://www.amia.org/pubs/symposia/D005602.PDF>.
- ERADUS, W.J. et al. Oestrus detection in dairy cattle using a fuzzy inference system. **Application and Ergonomics in Agriculture**, v.14, n.1, p.185-188, 1998.
- FAYYAD, U.; STOLORZ, P. Data mining and KDD: promise and challenges. **Future Generation of Computer Systems**, v.13, p.99-115, 1996.
- FRENCH, P.D.; NEBEL, R.L. The simulated economic cost of extended calving intervals in dairy herds and comparison of reproductive management programs. **Journal of Dairy Science**, v.86, supl.1, p.52-54, 2003.
- FIRK, C.A. et al. Automation of oestrus detection in dairy cows: a review. **Livestock Production Science**, v.75, p.219-232, 2002.
- HANSEN, D. **Vantagens e limitações das tecnologias de reprodução animal**. São Paulo: Lagoa da Serra, 2003. Capturado em: 10 out. 2004. Online. Disponível na Internet: http://www.lagoa.com.br/main_artigos.asp?idNoticia=2&opcao=Integra.
- HEERSCHKE, J.R.G.; NEBEL, R.L. Measuring efficiency and accuracy of detection of estrus. **Journal of Dairy Science**, v.77, n.9, 1994.
- KASTELIC, J.P. Computadorized heat detection. **Advancement in Dairy Technology**, v.13, p.393-402, 2001.
- KIDDY, C.A. Variation in physical activity as an indication of estrus in dairy cows. **Journal of Dairy Science**, v.60, p.235-243, 1977.
- NEBEL, R.L. et al. Automated electronic systems for the detection of estrus and timing of AI in cattle. **Animal Reproduction Science**, v.60, p.713-723, 2000.
- RORIE, R.W., et al. Application of electronic estrus detection technologies to reproductive management of cattle. **Theriogenology**, v.57, p.137-148, 2002.
- STURMAN, H., et al. Importance of inseminating only cow in estrus. **Theriogenology**, v.53, n.8, p.1657-1667, 2000.