International Journal of Livestock Research eISSN : 2277-1964 NAAS Score -5.36 | Vol 9 (01) Jan '19

Original Research

Application of Hidden Markov Models and Hidden Semi Markov Models as Diagnostic Aid in Chronic Mastitis in Cows

Safeer M. Saifudeen*, R. Venkataramanan and A. Serma Saravana Pandian

Department of Animal Husbandry Statistics and Computer Applications, Madras Veterinary College,
Tamil Nadu Veterinary and Animal Sciences University, Chennai, Tamil Nadu, INDIA

***Corresponding author:** safeermsaifudeen@gmail.com

Rec. Date:	Mar 02, 2018 04:33
Accept Date:	Oct 17, 2018 03:33
DOI	10.5455/ijlr.20180302043328

Abstract

Mastitis is the most important and expensive disease of dairy industry. The present study was conducted at the Large Animal Clinic of Madras Veterinary College Hospital, Chennai from November 2016 to March 2017 to find out the most probable sequence of stages of chronic mastitis. Analysis of HMM and HSMM were performed using 'The R Project for Statistical Computing' (version 3.4.0). The study summarized as disease started with stage of bacterial entry and its flare up (W1) with reduced milk production and presence of watery milk from the lactating quarters. As the time advanced, the disease progressed to the stage of increasing somatic cell count (W2) where repeated episodes of sub-acute mastitis could be seen. The stage of fibrous tissue proliferation is visible only in the case of hidden semi Markov model (HSMM). These models might be used in decision making process like, either maintenance or culling of the affected animals in established farms.

Key words: Chronic Mastitis, Inflammation, Hidden Markov Model, Hidden Semi Markov Model, Symptoms

How to cite: Saifudeen, S., Venkataramanan, R., & Pandian, A. (2019). Application of Hidden Markov Models and Hidden Semi Markov Models as Diagnostic Aid in Chronic Mastitis in Cows. International Journal of Livestock Research, 9(1), 53-60. doi: 10.5455/ijlr.20180302043328

Introduction

Mastitis is an important and costly disease of dairy sector which leads to huge economic losses to the farmers (Sharif and Muhammad, 2009). Those economic losses are due to reduction in production and productivity, medicine and labour costs and less market value. Clinical mastitis remains a complex disease and its control is a continue challenge despite of intensive research scheme (Bangar *et al*., 2016). In India pooled estimate of the prevalence of subclinical mastitis on cow-basis was 46.35 percentage (Bangar *et al*., 2015) and pooled estimates of clinical mastitis in crossbred cows on cow-basis was 16.08 percentage (Bangar *et al*., 2016). Mastitis is characterized by inflammation of mammary gland in response to injury for the purpose of destroying or neutralizing the infectious agents. Elevated leukocytes or somatic cells produced by inflammatory response cause a reduction in milk production and alter milk composition (Jones and Bailey, 2009). An inflammatory process that exists for months and may continue from one lactation to another is seen in chronic mastitis (Rabello *et al*., 2005). Viguier *et al*. (2009) stated that chronic mastitis is a rarer form of the disease, results in persistent inflammation of the mammary gland. There is high prevalence of subclinical mastitis in dairy cows in India, which might be responsible for low productivity in lactating cows in India over the years and needs to be controlled by adopting scientific, managemental and therapeutic measures. Dairy farmers can reduce incidence and economic losses due to subclinical mastitis under the guidance of field veterinarians (Bangar *et al*., 2015).

Hidden Markov model is a doubly stochastic process with an underlying stochastic process that is not observable (it is hidden), but can only be observed through another set of stochastic processes that produce the sequence of observed symbols (Rabiner and Juang, 1986). Hidden Markov models are important tools in estimation and analysis of biological sequences and many other systems (Pertsinidou and Limnios, 2015). One limitation of hidden Markov model (HMM) is that dwell times (time spent in a specific state) are assumed to follow a geometric distribution; this assumption may not always hold (Langrock *et al.*, 2012). The hidden semi- Markov model (HSMM), an extension of HMM, relaxes this assumption and permits explicit modeling of dwell times using alternative statistical distributions. Hidden Markov model has got wide application in veterinary field such as medicine, biochemistry, genetics, movement pattern of animals, vocal system, behavioural aspects etc. Though studies have been done in the disease progression modeling of Alzheimer's disease in humans (Sukkar *et al.*, 2012), such attempts are scarce in veterinary field. Considering the fact that a clear cut idea about the progression of chronic cases of mastitis in a herd is essential for effective mastitis prevention and control program, the present study has been designed with the objectives including tracing the unknown sequence of stages of chronic mastitis in cows and to find out the most probable sequence of stages of chronic mastitis along with its most recurring symptoms.

Materials and Methods

The present study was conducted at the Large Animal Clinic of Madras Veterinary College (MVC) Hospital, Chennai from November 2016 to March 2017. The primary data were collected from mastitis affected cows presented in outpatient ward of the MVC hospital. Out of 127 cases studied, 60 were diagnosed as chronic mastitis. Each and every case was observed thoroughly for stages of disease and possible symptoms within stage. The particular stage of the disease (which was hidden) and the change in stage of disease were identified by the symptoms shown by the affected animals. The probability of occurrence of each and every symptoms at particular stages and the probabilities of all possible transition of stages from one to another was required for finding out the most probable sequence of stages of mastitis.

DOI 10.5455/ijlr.20180302043328

For chronic cases of mastitis, the three stages of disease considered were stage of bacterial entry and its flare up (W_1), stage of increasing somatic cell count (W_2) and the localisation of infectious organism and stage of fibrous tissue proliferation (W_3). For each stage, symptoms observed (V_{xy}) were classified into three groups and was made uniform for all disease stages. Observed types of symptoms (in general) included were abnormal milk, repeated episodes of sub clinical mastitis and clinical signs.

The probability of transition from one stage to other stage or itself (transition probability) and the probability of showing particular symptoms at different stages (emission probability) were calculated. Out of total number of affected animals, the symptoms showed at each transition disease stage were noted. Probability of occurrence of each symptom at different stages were calculated by finding out the number of animals exhibited that particular symptom divided by total number of affected animals. The most probable symptom at each stage could be identified. Detailed description of hidden stages of mastitis, its transition probabilities, the observed symptoms along with the emission probabilities are given in Tables 1 and 2. Using the above said probability values both transition and emission probability matrices were constructed.

Computation of Hidden Markov Model

Decoding

The decoding problem could be used to find the most probable sequence of hidden states provided a sequence of visible states V^T. For this every possible path should be enumerated and calculate the probability of the visible sequence observed. Simple decoding algorithm could be used. Viterbi algorithm is most commonly used for decoding (Pertsinidou and Limnios, 2015, Lou, 1995). Algorithm works by doing iteration with all the possible sequence of stages. These algorithms can be used to decode an unobserved hidden semi-Markov process and it is the first time that the complexity is achieved to be the same as in the Viterbi for Hidden Markov models (Rabiner and Juang, 1986).

Statistical Analysis

Analysis of HMM and HSMM were performed using 'The R Project for Statistical Computing' (version 3.4.0). Package 'HMM' by scientific software development (Lin Himmelmann, 2010) was used as the model programme to compute the most probable path of states for a sequence of observations for a given Hidden Markov Model using Viterbi algorithm. Similarly, Package 'hsmm' (Jan Bulla *et al.*, 2010) using Viterbi algorithm was used in Hidden Semi Markov Model (HSMM). Using HSMM, duration of stay at each stage of mastitis were incorporated in addition to the above said information to get the better picture of the progression of chronic mastitis.

DOI 10.5455/ijlr.20180302043328

Result and Discussion

The transition probability and emission probability matrices in the present study were obtained as -

Transition probability matrix

$$\begin{bmatrix} 0.30 & 0.40 & 0.30 \\ 0.20 & 0.15 & 0.15 \\ 0.10 & 0.15 & 0.15 \end{bmatrix}$$

Emission probability matrix

$$\begin{bmatrix} 0.30 & 0.30 & 0.40 \\ 0.40 & 0.30 & 0.30 \\ 0.40 & 0.45 & 0.15 \end{bmatrix}$$

Using these limited informations, the hidden stages of mastitis could be found out using hidden Markov model. Similar observation were made by Martin *et al.* (2006). They concluded that the hidden Markov model achieved valuable prediction results using only a limited number of parameters. Similar observation were made by Doroshenkov *et al*. (2007). Transition and emission probability matrices produced were used to create the programme in 'The R Project for Statistical Computing' for computation of most probable path of states using Viterbi algorithm. Pertsinidou and Limnios (2015) presented a new Viterbi algorithm for Hidden semi-Markov models. The Viterbi algorithm could be used to decode an unobserved hidden semi-Markov process same as in the Viterbi for Hidden Markov models. Similarly, O'Connell and Hojsgaard (2011) described the R package for modified hidden semi Markov model (MHSMM) which implemented estimation and prediction methods for hidden Markov and semi-Markov models for multiple observation sequences Most probable hidden stages of mastitis were identified using the probability of observed symptoms and the probability of transition from one stage to another. Similar attempt were made by Ohlsson *et al.* (2001).

Stages of chronic mastitis along with its observed symptoms were given in Table 1. Probability of occurrence of each symptom at particular stages were calculated. The most probable symptom at each stage could be identified. All the possible probabilities of the transition of stages were given in Table 2. Pictorial representation of the probabilities were given in Figure 1 where 'W' represents stages of mastitis and 'V' represents observed symptoms of disease.

Table 1: Stages of mastitis with the observed symptoms

Labels for Hidden Observation	Observed Types of Symptoms in General	Probability of Occurrence Noted (in percentage)
Stage of bacterial entry and its flare up (W1)	Abnormal milk	91
	Repeated episodes of sub clinical mastitis	89
	Clinical signs	96
Stage of increasing somatic cell count (W2)	Abnormal milk	97
	Repeated episodes of sub clinical mastitis	91
	Clinical signs	91.5
Localisation of infectious organism and stage of fibrous tissue proliferation (W3)	Abnormal milk	88
	Repeated episodes of sub clinical mastitis	90.5
	Clinical signs	31

DOI 10.5455/ijlr.20180302043328

Table 2: Transition stages of mastitis along with its probability of occurrence

Labels for Hidden Observation	Possible Transitions	Probability of Occurrence Noted (in percentage)
Stage of bacterial entry and its flare up (W1)	W1 to W1	30
	W1 to W2	40
	W1 to W3	30
Stage of increasing somatic cell count (W2)	W2 to W1	20
	W2 to W2	15
	W2 to W3	15
Localisation of infectious organism and stage of fibrous tissue proliferation (W3)	W3 to W1	10
	W3 to W2	15
	W3 to W3	15

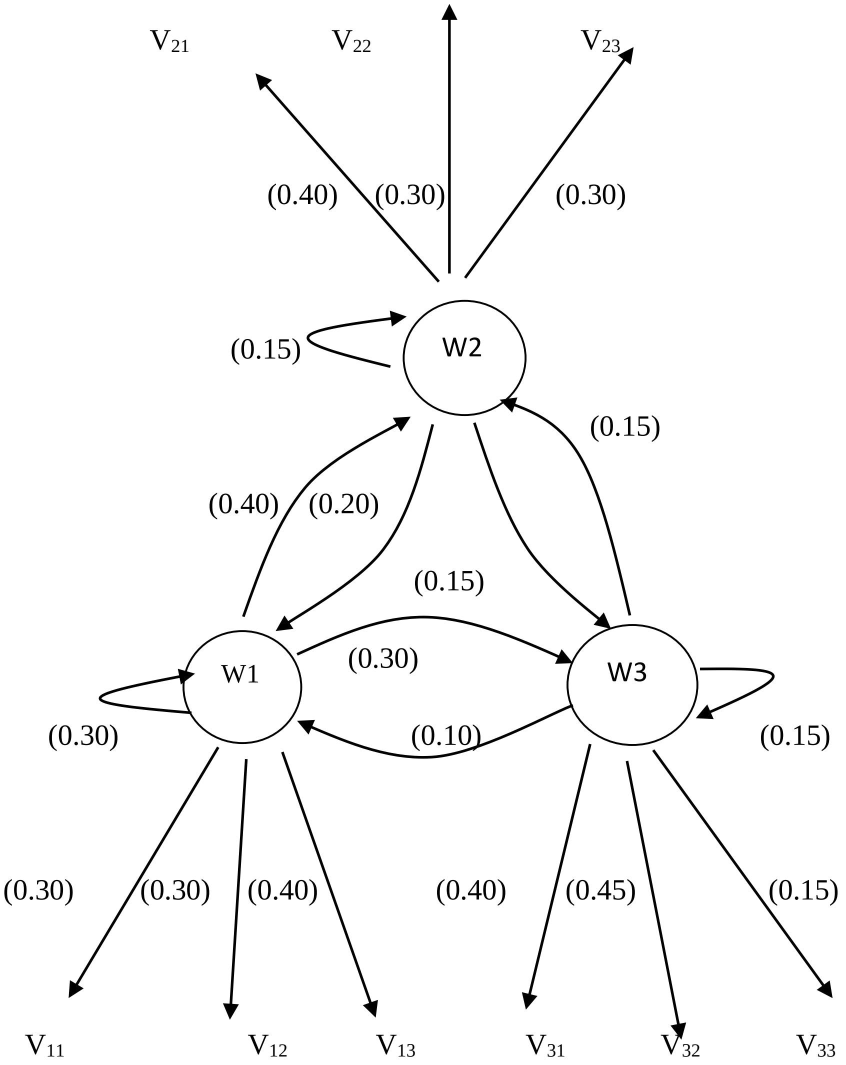

Fig.1: Stages of mastitis represented as three state ergodic Hidden Markov model. W-Hidden stages of mastitis: V-Symptoms produced by each stage (visual observations). Transition and emission probabilities were given in the parenthesis.

DOI 10.5455/ijlr.20180302043328

Three staged disease model with three prominent symptoms at each stage were considered in case of chronic mastitis. This is in conformity with earlier studies made by Silva (2003). He concluded that a three-state model provided a good description for the data considered.

Using hidden Semi Markov model, duration of stay at each stage of mastitis were incorporated in addition to the above said information to get the better picture of the progression of the disease. In the case of Hidden Semi Markov model, observed symptoms were expressed in terms of mean and variance. To convert the qualitative data into quantitative one, each observed symptom were assigned a score of 0.5. Mean and variance were found out for each stage of mastitis separately as given in Table 3.

Table 3: Mean and variance of observations for each stage in chronic mastitis

Labels for Hidden Observation	Observed Types of Symptoms in General	Mean of Observations*	Variance of Observations*
Stage of Bacterial entry and its flare up (W1)	Abnormal milk	1.6	0.11
	Repeated episodes of sub clinical mastitis		
	Clinical signs		
Stage of increasing somatic cell count (W2)	Abnormal milk	1.76	0.13
	Repeated episodes of sub clinical mastitis		
	Clinical signs		
Localisation of infectious organism and stage of fibrous tissue proliferation (W3)	Abnormal milk	0.87	0.09
	Repeated episodes of sub clinical mastitis		
	Clinical signs		

**score of 0.5 is given to each observed symptom*

As the stages of chronic diseases are differentiable and the time of stay in each stage is highly evident, hidden Markov model and hidden semi Markov model could be exploited in finding the progression of chronic diseases. Here in the case of chronic mastitis, results obtained revealed that only the first two stages, stage of bacterial entry and its flare up (W1) and stage of increasing somatic cell count (W2) were visible in the case of hidden Markov model. The results obtained are summarised in Table 4.

Table 4: Model summary

Model Used	Most Probable Sequence of Stages of Disease
Hidden Markov model	*"W1" "W2" "W1" "W2" "W1" "W1" "W1" "W1" "W1"*
Hidden semi Markov model	*W1W1W1W1W1W1W1W1 W3 W3W3 W3 W2 W2W2*

The overall disease progression could be concluded as follows. The disease started with stage of bacterial entry and its flare up (W1) with reduced milk production and presence of watery milk from the lactating quarters. As the time advanced, the disease progressed to the stage of increasing somatic cell count (W2) where repeated episodes of sub-acute mastitis could be seen. In hidden Markov model (HMM), most of the animals retained in stage one with reduced milk production, reduced milk quality and the presence of watery

DOI 10.5455/ijlr.20180302043328

milk from the lactating quarters. In hidden semi Markov model after a remarkable period of stay in the first stage of disease, the disease stage progressed to stage three (W3) where local areas of fibrosis were evident in udder with poor response of the animals to treatment procedures. Atrophy of gland was also noted in 31 percent of cases.

The stage of fibrous tissue proliferation (W3) is visible only in the case of hidden semi Markov model (HSMM). Even though probability of occurrence of the third stage (W3) was less when compared to other stages, the time of stay in W3 is more. Once the localisation of infectious organism and the stage of fibrous tissue proliferation occurred, recovery chance would be less or the time taken for complete recovery would be more. These points were clearly demarcated by hidden Semi Markov model. In agreement with the current study, Luo *et al*. (2013) developed an approach for estimating chronic kidney disease stage transition rates using hidden semi Markov models (HSMMs), when the level of information and observation time varied among individuals. A hidden semi-Markov model was proposed by Altuve *et al.* (2011) in apnea-bradycardia detection to represent and characterize the temporal evolution of observed RR series and different pre-processing methods of these series were investigated. They compared two detectors of apnea-bradycardia episodes, based on hidden Markov models (HMM) and hidden semi-Markov models (HSMM).

Conclusion

The present study revealed the unknown sequence of stages of chronic mastitis in cows and the most probable sequence of stages of chronic mastitis along with its most recurring symptoms were found out. In chronic mastitis, results obtained revealed that only the first two stages, stage of bacterial entry and its flare up (W1) and stage of increasing somatic cell count (W2) were visible in the case of hidden Markov model. The stage of fibrous tissue proliferation (W3) is visible only in the case of hidden semi Markov model (HSMM). Even though probability of occurrence of the third stage (W3) was less when compared to other stages, the time of stay in W3 is more. Hidden Markov Model (HMM) and Hidden Semi-Markov Model (HSMM) might be explored for finding the most probable stages of other relevant diseases especially the chronic diseases. Further these models might be used in decision making process like, either maintenance or culling of the affected animals in established farms. Similar studies might be conducted in veterinary dispensaries and clinics and the results obtained might be compared.

Acknowledgement

The authors are thankful to Dean and to Head of Department of Clinics of Madras Veterinary College for providing the necessary facilities to conduct the research work. Authors extend their sincere thanks to employees of college hospital including duty doctors, staffs for their co-operation during data collection.

References

1. Altuve, M., G. Carrault, A. Beuchee, P. Pladys and A. I Hernandez (2011). On-line apnea-bradycardia detection using hidden semi-Markov models. *In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*: pp. 4374-4377.
2. Bangar, Y.C., B. Singh, A. K. Dohare and M. R. Verma (2015). A systematic review and meta-analysis of the prevalence of subclinical mastitis in dairy cows in India. *Tropical Animal Health & Production*, 47, 291–297.
3. Bangar, Y.C., M. R. Verma, A. K. Dohare and R. Mukherjee (2016). Meta-analysis of the prevalence of clinical mastitis in crossbred cows in India. *Journal of Animal Research*, 6(6), 933-938.
4. Doroshenkov, L. G., V. A. Konyshev and S. V. Selishchev (2007). Classification of human sleep stages based on EEG processing using hidden Markov models. *Biomedical Engineering*, 41(1): 25-28.
5. Himmelmann, L (2010). HMM – Hidden Markov Models. R Package Version 1.0, URL http://CRAN.R-project.org/package=HMM.
6. Jan Bulla, Ingo Bulla and Oleg Nenadic (2010). hsmm - An R package for analyzing hidden semi-Markov models. Computational Statistics & Data Analysis. 54: 611-619.
7. Jones, G. M. and T. L. Bailey (2009). Understanding the basics of mastitis. *Virginia Cooperative Extension*, 404: 233.
8. Langrock, R., R. King, J. Matthiopoulos, L. Thomas, D. Fortin and J. M. Morales (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, 93(11): 2336-2342.
9. Lou, H. L. (1995). Implementing the Viterbi algorithm. IEEE signal processing magazine, 12(5): 42-52.
10. Luo, L., D. Small, W. F Stewart and J. A. Roy (2013). Methods for estimating kidney disease stage transition probabilities using electronic medical records. *Generating Evidence & Methods to improve patient outcomes*, 1(3): 6.
11. Martin, J., J. F Gibrat and F. Rodolphe (2006). Analysis of an optimal hidden Markov model for secondary structure prediction. *BMC Structural Biology*, 6(1): 25.
12. O'Connell, J., S. Højsgaard (2011). Hidden Semi Markov models for multiple observation sequences: *The mhsmm package for Journal of Statistical Software*, 39(4): 1-22.
13. Ohlsson, M., C. Peterson and M. Dictor (2001). Using hidden Markov models to characterize disease trajectories. *In Proceeding of the neural networks and expert systems in medicine and healthcare conference, Washington* pp. 324-326.
14. Pertsinidou, C. E. and N. Limnios (2015). Viterbi algorithms for Hidden semi-Markov Models with application to DNA Analysis. *RAIRO- Operations Research*, 49(3): 511-526.
15. Rabello, R. F., C. R. V. M. Souza, R. S. Duarte, R. M. M. Lopes, L. M. Teixeira and A. C. D. Castro (2005). Characterization of *Staphylococcus aureus* isolates recovered from bovine mastitis in Rio de Janeiro, Brazil. *Journal of Dairy Science*, 88(9): 3211-3219.
16. Rabiner, L., B. Juang (1986). An introduction to hidden Markov models. *IEEE ASSP magazine*, 3(1): 4-16.
17. Sadashiv, S. O. and B. Kaliwal (2013). Prevalence of bovine mastitis in north Karnataka, India. *International Journal of Pharma and Health Care Research*. 01, 169-177.
18. Sharif, A., G. Muhammad (2009). Mastitis control in dairy animals. *Pakistan Veterinary Journal*, 29(3): 145-148.
19. Silva, C. Q. D (2003). Hidden Markov models applied to a subsequence of the Xylella fastidiosa genome. *Genetics and Molecular Biology*, 26(4): 529-535.
20. Sukkar, R., E. Katz, Y. Zhang, D. Raunig and B. T. Wyman (2012). Disease progression modelling using hidden Markov models. In Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, London 23-7-12 pp. 2845-2848.
21. Viguier, C., S. Arora, N. Gilmartin, K. Welbeck and R. O'Kennedy (2009). Mastitis detection: Current trends and future perspectives. *Trends in Biotechnology*. 27(8): 486-493.

DOI 10.5455/ijlr.20180302043328