

# The importance of ‘memory’ in statistical models for animal feeding behaviour

David J. Allcroft<sup>1</sup>, Bert J. Tolkamp<sup>2</sup>, Chris A. Glasbey<sup>1</sup> and Ilias Kyriazakis<sup>2</sup>

<sup>1</sup> *Biomathematics & Statistics Scotland, King’s Buildings, Edinburgh, EH9 3JZ, Scotland*

<sup>2</sup> *Animal Nutrition and Health Department, SAC, King’s Buildings, Edinburgh, EH9 3JG, Scotland*

Corresponding author: David Allcroft, email: [dave@bioass.ac.uk](mailto:dave@bioass.ac.uk), tel: 0131 650 8740, fax: 0131 650 4901

We investigate models for animal feeding behaviour, with the aim of improving understanding of how animals organise their behaviour in the short term. We consider three classes of model: latent Gaussian, hidden Markov and semi-Markov. Each can predict the typical ‘clustered’ feeding behaviour that is generally observed, however they differ in the extent to which ‘memory’ of previous behaviour is allowed to affect future behaviour. The hidden Markov model has ‘lack of memory’, the current behavioural state being dependent on the previous state only. The latent Gaussian model assumes feeding/non-feeding periods to occur by the thresholding of an underlying continuous variable, incorporating some ‘short-term memory’. The semi-Markov model, by taking into account the duration of time spent in the previous state, can be said to incorporate ‘longer-term memory’. We fit each of these models to a dataset of cow feeding behaviour. We find the semi-Markov model (longer-term memory) to have the best fit to the data and the hidden Markov (lack of memory) model the worst. We argue that in view of effects of satiety on short-term feeding behaviour of animal species in general, biologically suitable models should allow ‘memory’ to play a role. We conclude that our findings are equally relevant for the analysis of other types of short-term behaviour that are governed by satiety-like principles.

**Keywords:** Feeding behaviour; Hidden Markov model; Latent Gaussian model; Semi-Markov model; Short-term behaviour; Underlying states.

## 1 INTRODUCTION

Animal behaviour data can be collected in many ways and for a variety of reasons. Interest sometimes centres on long-term behaviour, looking at the overall behavioural profiles of animals and at whether specific goals have been achieved. Sufficiently long periods of time must be considered and interest lies mainly in summaries of the data. For feeding behaviour, such summaries might be the mean proportion of time feeding, the average frequency of feeding events or, if more than one food is available, the average diet selected (e.g. Kyriazakis and Emmans, 1991; Friggens et al., 1998; Tolkamp et al., 1998). Although these summaries might be adequate for certain purposes, essential information about the short-term organisation of feeding behaviour will be lost. But this is exactly our interest here: how are individual behavioural events organised and how does this relate to the underlying biology? By consideration of

Characteristic	Mean	SD	Range
Live weight (kg)	632	61	546–729
Daily milk yield (kg)	33.1	5.7	26.5–45.3
Daily fresh food intake (kg)*	46.3	3.8	40.3–54.0
No. of daily visits	26.8	8.0	16.8–44.1
Total daily feeding time (min)	147	31	108–217
Median visit duration (min)	4.1	0.7	3.2–5.5

Table 1: Background information regarding the data set: mean and variation in individual animal performance during the observation period.

\* average dry matter content of the food was 465 g/kg of fresh food

potentially suitable models, we hope to gain some insight into how short-term behaviour affects long-term performance (e.g. Tolkamp et al., 2002). A good model should not only fit the data well from a statistical point of view, but also provide an interpretation that is biologically sound.

For the purpose of this paper, we fit three classes of model to a set of short-term cow feeding data. The dataset is described first, the models subsequently. The models are (i) a latent Gaussian variable model, (ii) a hidden Markov model and (iii) a semi-Markov model. All these models can predict the typical clustered feeding behaviour that is frequently observed, but they differ in the extent to which memory of previous behaviour is allowed to affect future behaviour. First we determine how these models perform statistically. Subsequently, we analyse how well the underlying assumptions of the models agree with biological principles. We present our findings in the light of observations of short-term behaviour in other species and discuss how applicable our conclusions are for animal behaviour in general.

## 2 DATA

The data consist of complete records of feeding behaviour for eight cows, recorded by a set of computerised feeders at the Langhill Dairy Cattle Research Centre, Roslin, Midlothian during a 30-day period in April–May 1995. Food, consisting of a mixture of grass silage and concentrates, was supplied to a group of cows, all of which were multiparous lactating Holstein-Friesians. The cows were in a single yard and, except for short periods during milking and fresh food supply, had continuous access to the feeders. All animals were allowed to adapt to the food, feeding system and management procedures for a period of at least six weeks before the data were collected. Transponders, worn around their necks, gave them access to six feeders that supplied their particular food. The start and stop times of all visits were recorded to the nearest second and these records formed the basis of our analysis. Table 1 provides some background information on animal performance during the measurement period. Tolkamp and Kyriazakis (1997) and Tolkamp et al. (1998) provide full details about the experimental facility, management of the animals and composition of the food.

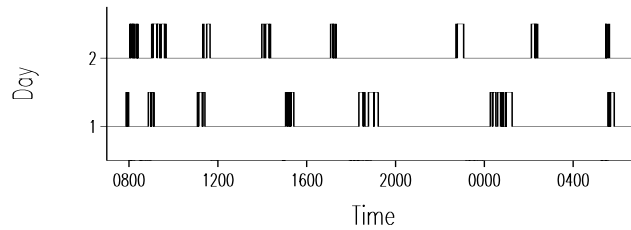


Figure 1: Two days of feeding data for one of the cows. Raised values of the signal denote periods of feeding.

An example of the visit structure for one of the animals is given in Figure 1. The cows typically ate about two-thirds of their food during the day (0800 to 2000 hours) and the rest during the night (Tolkamp et al., 2002). Inspection of feeding patterns for all cows showed there to be large individual variation in feeding patterns, with some animals being very regular in their feeding times, and others being much less consistent. The best modelling approach therefore is to fit models to individual animals and then summarise parameters over the group.

The data obtained were recorded in continuous time (to the nearest second), however both the hidden Markov and latent Gaussian models can only be formulated in discrete time. Therefore for these models, the data must first be discretised. We chose to do this on a 1-minute scale, i.e. if during a given minute the cow was visiting a feeder for more than 30 seconds, the feeding variable took the value 1, otherwise the feeding variable was equal to 0. The 1-minute time scale is arbitrary, being a compromise between retaining enough detail in the data and dealing with series of manageable length (for each cow we used 30 days corresponding to a time series of 43200 minutes of observations).

### 3 MODELS

A simple inspection of the data leads us to conclude that the general structure of the feeding behaviour can be summarised as in Figure 2. Feeding events separated by short gaps are clustered into bouts; these bouts are then separated from each other by longer gaps. So even during the short non-feeding periods within a bout, the animal can still be considered to be within an overall feeding state. These short gaps within a bout could be said to be a consequence of the nature of feeding behaviour in our experimental facility, i.e. an animal moving from one feeder to another. However, similar clustered feeding behaviour has frequently been observed in cows (e.g. Metz, 1975; Dado and Allen, 1993; Rook and Huckle, 1997; Stamer et al., 1997), as well as other species such as horses (Mayes and Duncan, 1986), pigs (Morgan et al., 2000), rats (Collier et al., 1999), starlings (Langton et al., 1995), zebra finches (Sibly et al., 1990) and blowflies (Simpson et al., 1989). This suggests that a model with some sort of ‘latent’ structure could

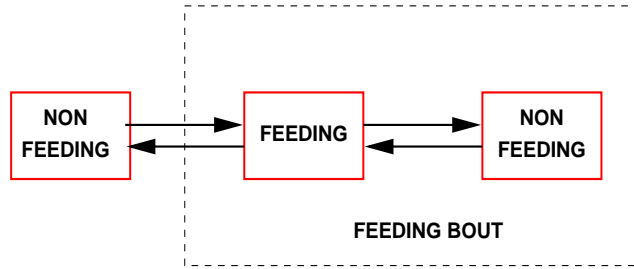


Figure 2: Diagrammatic representation of the possible model structure.

be widely appropriate for modelling feeding behaviour; this would allow for changes in observed behaviour even when there is no change in the overall behavioural state of the animal. The main difference between the proposed models is that they incorporate this idea of latent structure in different ways. The latent Gaussian model has a continuously varying latent variable, whilst the hidden Markov model has a latent sequence of discrete states. The semi-Markov model has no explicit latent component but one can be superimposed after fitting, making this a potentially suitable model too. We will discuss these issues and other details of model fitting for each model separately. Mathematical details have been kept to a minimum, but references are provided where appropriate.

#### (a) Hidden Markov model

Markov models are commonly used to analyse time sequences of data in which an animal moves between states according to a set of transition probabilities (see, for example, Haccou and Meelis, 1994). A hidden Markov model (HMM), described fully by MacDonald and Zucchini (1997), is a model for which a series of unobserved underlying states follow a Markov chain, the observed behaviours then being conditionally independent of each other and dependent only on the current underlying state. MacDonald and Raubenheimer (1995) used these models to study the perambulatory behaviour of locusts, suggesting that in terms of model parsimony, a HMM can be a better alternative to a higher order Markov model in cases where a first order Markov model is inadequate. HMMs can only be formulated in discrete time, nevertheless they are still biologically attractive because the underlying state of the animal can be modelled rather than the observed behaviour. The most obvious HMM for feeding data with a structure as depicted in Figure 2 would have two states. The first of these would be a ‘non-feeding’ state, corresponding to the long intervals between bouts. The second would be a ‘feeding’ state which would allow for the whole feeding bout, i.e. both the feeding events themselves and the short non-feeding intervals that separate them when the animal is moving between feeders. The hidden Markov model explicitly allows for this change in observed behaviour without any change in the underlying state of the animal. Part of an example realisation is presented in Figure 3, showing how an animal can remain within a feeding bout whilst displaying short periods of non-feeding behaviour, resulting in clusters of

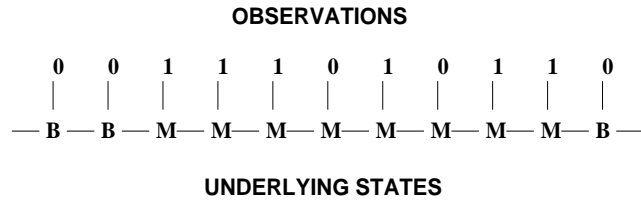


Figure 3: Realisation of a hidden Markov model. The underlying state (**M**eal or **B**etween-meal) follows a Markov chain. The probability of the observed behaviour being feeding (**1**) or non-feeding (**0**) are dependent on the current underlying state, i.e. **B** or **M**.

feeding events as observed in the data.

In terms of model fitting, two sets of parameters have to be estimated: (i) a set of transition probabilities for moving between the underlying states, and (ii) a probability of feeding within each state. The likelihood function can be written down explicitly, as a matrix product of transition probabilities and state-dependent feeding probabilities (see, for example, MacDonald and Zucchini, 1997, Section 2.5), and parameters can be estimated by maximum likelihood using a numerical optimisation routine. Diurnal patterns can be allowed for, either at the level of the transition probabilities or the state-dependent probabilities of feeding. In either case, the constant probabilities are replaced with functions of time of day (see MacDonald and Zucchini, 1997, for full details). It should be emphasised that this is exclusively a time of day effect and nothing to do with the length of time in the current or previous states, as is the case for the semi-Markov models considered later.

For a HMM with two states, two parameters for the transition matrix ( $\Gamma$ ) have to be estimated. These represent the probabilities of animals remaining in (and, by implication, the probabilities of animals leaving) the current state at the next time-step (i.e. minute). The following matrix contains parameter estimates for one of the animals:

$$\Gamma = \begin{pmatrix} .9904 & .0096 \\ .0678 & .9322 \end{pmatrix}.$$

Hence the probability of this animal remaining in a non-feeding state from one minute to the next is 0.9904, whilst the probability of it remaining in a feeding state is 0.9322 (the other entries are determined by the rows being constrained to sum to 1). For the same animal, the state-dependent probabilities of feeding were estimated as  $p = (.0000, .9604)$ , indicating a zero probability of feeding in the non-feeding (between bout) state and a 0.9604 probability of feeding in the feeding (bout) state. The transition matrix  $\Gamma$  also determines the overall probabilities of being in the non-feeding or feeding states as  $(.8757, .1243)$ . Hence in this model, the animal spends about 88% of its time in a non-feeding state, and the remaining 12% in a feeding state within which it spends 96% of its time feeding and the other 4% not feeding. This demonstrates how the hidden Markov model is able to include short non-feeding periods within the feeding state.

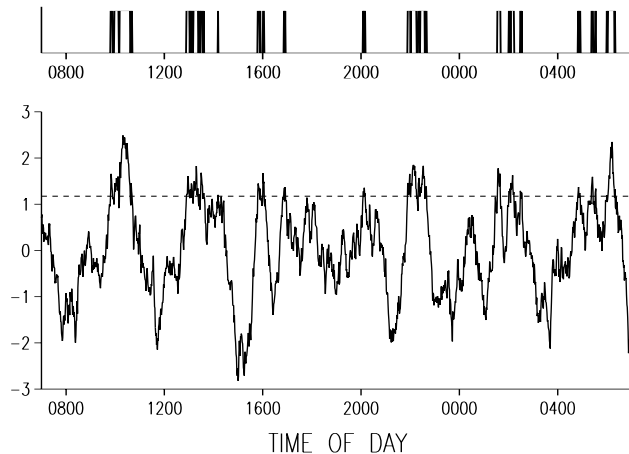


Figure 4: Simulations from a latent Gaussian model, fit to one of the cows. The threshold level is chosen to match the observed rate of feeding and the resulting feeding events are shown above.

### (b) Latent Gaussian model

This model is based on the assumption that the binary feeding/non-feeding data arise from the thresholding of a continuous, normally distributed variable, such that periods of feeding correspond to the variable exceeding some threshold. Biologically, the idea is that the latent variable corresponds to some physiological or neurological state of the animal that affects its motivation to feed. Figure 4 shows simulations of such a variable, using parameter values estimated for one of the animals. Although the model is fit in discrete time (1-minute scale), the level of the latent variable can be considered to change continuously, and the crossing of the threshold corresponds to the animal being motivated to either resume or stop feeding. The resulting pattern of feeding events in Figure 4 shows that this model can indeed give rise to data similar to the observed (e.g. Figure 1).

It is important to appreciate that the observed binary series can be considered a censored version of the unobserved Gaussian series. For a given binary series of feeding, we cannot recover the actual trace of the Gaussian series as shown in Figure 4, however we can make use of the one-to-one correspondence that exists between the auto-correlation of the two series. Hence from the sample auto-correlation of the binary series at a given time lag we can estimate the auto-correlation of the Gaussian series at that lag (Allcroft and Glasbey, 2002b). The set of auto-correlation coefficients at different time lags then characterises the Gaussian series completely and it is not necessary to recover the actual trace of the series. The threshold level, shown in Figure 4, must also be estimated for each individual. This value is fixed at the level for which the proportion of time spent over the threshold matches the observed proportion of the overall time spent feeding. Here it is also possible to take diurnal effects (or other covariates) into account, by general-

ising the constant threshold to one that varies with time of day. However this results in very little change in the auto-correlation estimates (Allcroft, 2001) and so for simplicity we ignore this here. The next step is to fit a model to the estimated auto-correlation structure. A convenient class of models to consider is that of auto-regressive moving-average (ARMA) processes (see, for example, Box et al., 1994). These are an established class of statistical models for Gaussian time series that exhibit short-term memory. The simplest class of ARMA model that can produce the observed shape of auto-correlation function is of order (2,1), i.e. there are two auto-regressive parameters  $\phi = (\phi_1, \phi_2)$  and one moving-average parameter  $\theta$ . This roughly corresponds to the value of the continuous variable being dependent on the observed value of the variable at the previous two time-steps and on the expected value at the previous time-step, given what went before. A consequence of dealing with censored data is that full maximum likelihood cannot be used for parameter estimation. Instead, we simply used least squares to match the estimated auto-correlation coefficients to their expected values under the ARMA(2,1) model. More details about this and other model fitting procedures are given by Allcroft and Glasbey (2002b).

### (c) Semi-Markov model

Markov processes can generally be formulated in both discrete and continuous time. For a model in continuous time, if the duration of each event is ignored, the sequence of states follows a discrete Markov chain (Haccou and Meelis, 1994). This is also the case for a semi-Markov model. The Markov, or lack of memory, property is that the probability of transition to the next state is dependent only on the current state and not on past states. A consequence of this property is that the distributions of durations in each state have exponential distributions (Haccou and Meelis, 1994). A semi-Markov model relaxes this property and, as a consequence, event durations can follow other specified distributions, in general dependent on the event type, e.g. feeding/non-feeding.

If the probability of an animal starting to feed was constant (i.e. irrespective of the time since it fed last), feeding events would occur randomly in time and durations between feeding events would be exponentially distributed. When feeding events are separated by short non-feeding periods and bouts of feeding events are separated from other bouts by longer non-feeding periods, we might expect the overall population of non-feeding durations to be a mixture of two exponentially distributed populations (e.g. Slater and Lester, 1982). However, previous analysis of cow feeding data showed that the marginal distribution of durations between feeding events could be better described by a mixture of two log-normal distributions (Tolkamp et al., 1998). This is also the case for the data we analyse here, Figure 5 showing the fit for one of the cows. Hence non-feeding events can be so classified as within or between bouts (Tolkamp et al., 1998). Tolkamp and Kyriazakis (1999) investigated the occurrence of a third distribution for some animals, caused by drinking behaviour. However for simplicity, and to remain more comparable with the other two types of model, we consider only mixtures of two distributions here. The semi-Markov model we fit is therefore based upon the following four considerations. Firstly, the marginal distribution of

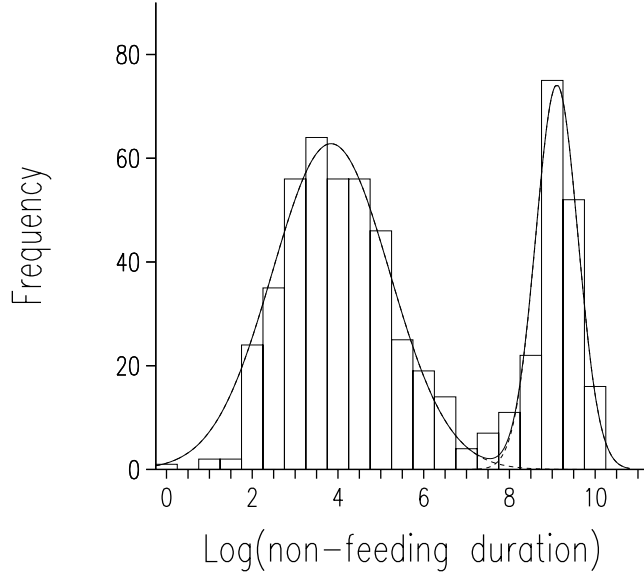


Figure 5: Distribution of log-transformed non-feeding durations, with fitted mixture of two log-normal distributions.

visit durations is adequately described by an exponential distribution (Tolkamp et al., 2000). Secondly, a mixture of two log-normal distributions is appropriate to describe the marginal distribution of non-feeding durations (Tolkamp et al., 1998). Thirdly, the durations of adjacent feeding events show no obvious dependency and are therefore assumed independent. And, lastly, there is some dependence between adjacent non-feeding events when classified as short (within bout) or long (between bouts). For full justification of these considerations, see Allcroft (2001). For feeding states, only one parameter (for the exponential distribution of durations) has to be estimated. Non-feeding periods are not always obviously within a bout or between bouts, therefore the type must be treated as missing, and we can use essentially the same methodology as that used to fit the HMMs of the previous section (full details in this context are given in Allcroft, 2001). This results in simultaneous estimates of the two parameters for each log-normal distribution of within- and between-bout non-feeding durations, along with transition probabilities for whether the next non-feeding period is within or between bouts, given the type of the current one. Parameter estimates obtained for the eight cows are summarised in Table 2.

Although this model has no explicit latent structure, we can calculate a meal criterion from the mixture of log-normal distributions fit to the non-feeding durations. This meal criterion estimates the longest non-feeding period that can still be considered as within a bout, and corresponds to the duration at which the two log-normal distributions cross in Figure 5 (Tolkamp et al., 1998). This can then be used to group feeding events into meals and hence a pseudo-latent structure can be imposed on the model. Animals

Parameter	Mean	SD
Mean feeding duration (mins)	5.66	0.829
Mean within-meal non-feeding duration (mins)	2.66	1.018
SD within-meal non-feeding duration (mins)	4.14	2.180
Mean between-meal non-feeding duration (mins)	211	39.7
SD between-meal non-feeding duration (mins)	150	25.6
Overall proportion of non-feeding events that are between-meal	0.237	0.0639
Prob(next is within-meal given previous was within-meal)	0.719	0.0822
Prob(next is between-meal given previous was between-meal)	0.103	0.0568

Table 2: Summary of parameter estimates for the semi-Markov models fit to the eight cows.

can then be considered to move between a ‘feeding’ state (that includes the non-feeding gaps within a bout) and a ‘non-feeding’ (between-bout) state, in a similar way as with the HMM.

#### (d) Model comparison

Summarising the model descriptions, we observe that the dependence structure in each of the models is important in their initial motivation. The most important feature of the data is perhaps that the feeding events occur in bouts and not randomly in time. This in itself implies that there must be serial dependence in the data, otherwise the marginal distribution of the duration of non-feeding events would be well-described by a single exponential distribution (Metz, 1974). The three models considered allow for this in different ways but, biologically, all could be potentially plausible mechanisms for capturing data that show dependence.

Figure 6 shows the observed marginal distributions of the durations of feeding and non-feeding events, along with predictions from each of the models. The Y-axes are presented after square root transformation of the frequencies to allow easy inspection of the whole distribution on a single graph. For the feeding events (Figure 6a), all three models predict similar shapes, though visual inspection shows the latent Gaussian model to give the worst fit and the semi-Markov model the best. Figure 6b shows that the hidden Markov model is unable to capture the bimodal nature of the distribution of non-feeding durations and, as a consequence, many non-feeding periods which clearly occur within clusters of visits, have to instead be accounted for by the between-bout state. As a consequence, the model allowed only a very limited proportion of time within a bout to consist of non-feeding intervals (6%; see Table 3). Table 3 also shows that, as a direct result of assigning many shorter non-feeding periods to be between bouts, animals are estimated to have a large daily number of meals with short durations. It might be thought that the addition of a third state to the HMM might improve the fit, however the attractiveness of the hidden Markov approach is then lost, as all the three ‘states’ shown in Figure 2 would have to be included as underlying states and the hidden structure would be redundant. In addition, no matter how many extra states were added, the functional form of the distribution of non-feeding events would always be monotonic decreasing, hence the bimodal form of the distribution could still not be captured. Both the

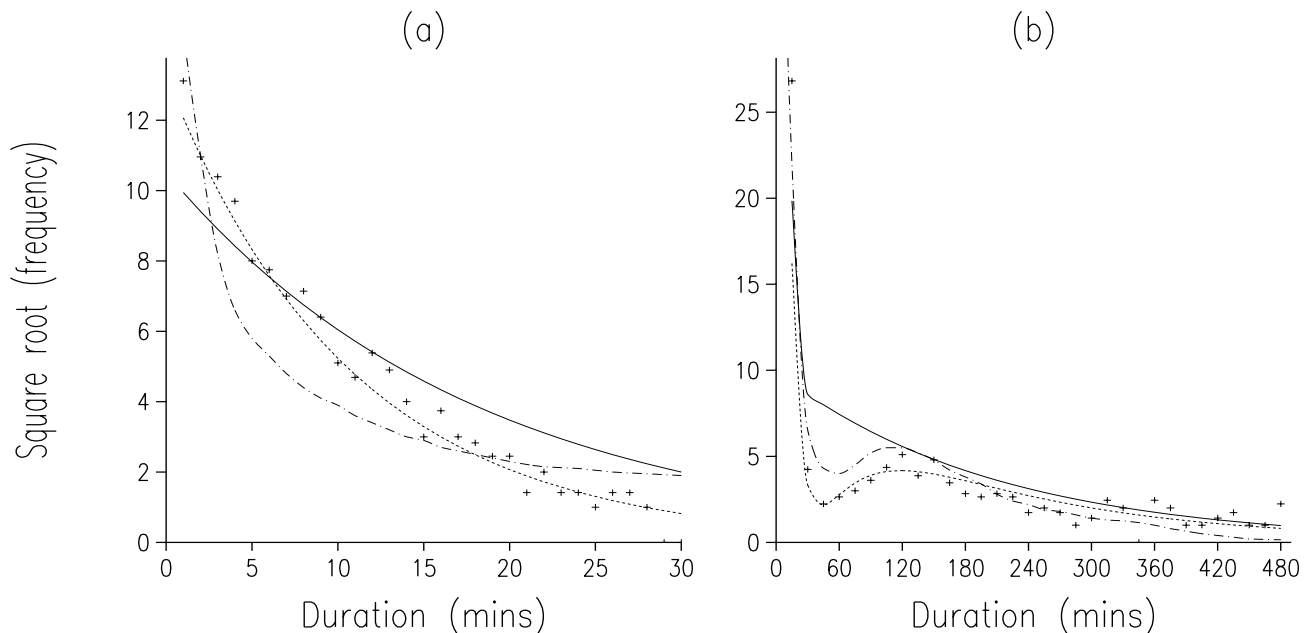


Figure 6: Marginal distributions for durations of (a) feeding events, (b) non-feeding events. (+) data; (—) hidden Markov model; (· — ·) latent Gaussian model; (· · ·) semi-Markov model.

Statistic	Hidden Markov		Semi-Markov	
	Mean	SD	Mean	SD
Meal duration (mins)	14.1	1.92	33.9	9.34
Between-meal duration (mins)	117	19.5	211	39.7
Daily number of meals	11.2	1.49	6.0	1.02
Overall proportion of time in ‘meal’ state	0.11	0.026	0.14	0.030
Proportion of time feeding within a meal	0.94	0.017	0.76	0.051
Overall proportion of time feeding	0.10	0.023	0.10	0.023

Table 3: Characteristics of meals as predicted by two of the models.

latent Gaussian and semi-Markov models do predict bimodal distributions of non-feeding durations similar to those observed (Figure 6b). It is not surprising that the semi-Markov model offers the slightly better fit, since the bimodal distribution is explicit in the motivation for this model. For the latent Gaussian model, the form of the marginal distributions played no part in its formulation and so it is reassuring that the shape of the bimodal distribution is predicted nevertheless. For this model, the plotted distributions were obtained by simulation from the fitted model, as analytic forms are not available. As no natural definition of a meal results from the latent Gaussian model, this model is not included in Table 3. With the semi-Markov model, quoting average values, we have six meals per day, each lasting about 30 minutes, separated by between-bout gaps averaging 3.5 hours. Within meals, about three-quarters of the time was spent feeding. This matches the observed data (e.g. Figure 1) much more than the HMM discussed

above, as now the within-meal gaps are being fully incorporated into the latent feeding state, as originally intended (Figure 2).

Thus far we have seen indications that the semi-Markov model is to be preferred over the other two, however no formal statistical comparison has been reported. For this we turned to a simulation-based method, involving the simulation of datasets from each fitted model and the subsequent re-fitting of each model to each set of simulated data. It is then possible to investigate which model the data most closely resemble. Full details of this procedure are not given here, but can be found in Allcroft and Glasbey (2002a). The analysis there showed that for five of the eight cows, there was no evidence that the data are inconsistent with the semi-Markov model, and for the other three cows, the semi-Markov model was preferred to the other two candidate models, although there was some evidence of lack of fit. Hence we conclude that, of the models considered, the semi-Markov model offers the best fit and, for most animals, is a good description.

## 4 DISCUSSION

We have considered three candidates for the modelling of cow feeding data, each with a potentially plausible biological motivation. Each of these models falls within the envisaged structure (Figure 2) and can produce results consistent with the clustering feeding behaviour observed in our data (Figure 1). Yet, our analysis reveals considerable differences in the descriptive power of these models.

A basic assumption in many Markov-type models is that event durations as well as inter-event times are distributed as a consequence of its ‘lack of memory’ property. However, in our data we found that only visit duration had an approximately exponential distribution (Figure 6a; see also Tolkamp et al., 2000, 2002). This would suggest that visits do indeed end randomly (Metz, 1974). However, if the Markov property applied to between-bout durations, bouts must occur randomly in time and then the distribution of non-feeding durations must also have an exponential distribution. Such distributions have often been assumed in analyses of feeding behaviour (e.g. Slater and Lester, 1982; Sibly et al., 1990; Langton et al., 1995). In addition, (hidden) Markov models have been used before in analyses of animal behaviour (MacDonald and Raubenheimer, 1995). However, an exponential distribution of durations between bouts would include a high frequency of short-duration periods (i.e. only one or a few minutes) as part of this distribution. This is the reason that the HMM allowed only very limited non-feeding time within bouts, as it classified many shorter non-feeding periods, that obviously occurred within clusters of visits, as between bouts. However, a strong feature of the data was that the distribution of inter-bout durations resembled a log-normal, with a very low frequency of short-duration periods (Figure 6b). Hence exponential distributions are not suitable for either population of durations between feeding events and simple Markov models are not appropriate.

The latent Gaussian model allows for effects of ‘short-term memory’. This in fact means that the probability of passing the threshold value at the next time-step not only depends on the current state of the animal but is also affected by the animal’s state at immediately previous times. This incorporation of some short-term memory results in a distribution of non-feeding durations close to that observed. However, as indicated by Figure 4, the underlying pattern is quite erratic in the sense that an animal in the non-feeding state may be ‘travelling’ towards the threshold value for some period of time and then reverse again and move away from the threshold without having consumed any food. Such behaviour contrasts with the way satiety is thought to affect the likelihood of animals initiating and terminating feeding bouts. This suggests that, for this dataset at least, a sequence of discrete latent states forms a better biological description of behaviour than a continuously varying latent variable.

The semi-Markov model came out as the most suitable overall, even though it has no explicit latent structure. However, such a structure is easily superimposed by prescribing that the animal should remain in a ‘bout state’ when it switches between the feeding state and a non-feeding period shorter than the meal criterion. Certainly one of the most important reasons for the good fit was that this model explicitly allows for the observed shape of the distribution of non-feeding periods. The log-normal model for the distribution of inter-bout durations is based upon the idea that during a feeding bout satiety will increase and that satiation will finally result in a feeding bout ending (Tolkamp et al. 2000). It seems reasonable to assume that, as a result of satiety, animals are unlikely to start another bout shortly after finishing one, hence the very low frequency of short non-feeding periods in this distribution (Tolkamp et al., 1998). As time passes, however, the likelihood of the animal starting a meal will continue to increase as satiety declines (Metz, 1975; Simpson, 1990; Tolkamp et al., 1998; Yeates et al., 2001). This is in strong contrast with the constant likelihood implicit in the HMM and also with the continuously ‘switching’ variable at the basis of the latent Gaussian model. The mixture of two log-normal distributions for non-feeding durations gives satisfactory results in our analysis and is consistent with the structure depicted in Figure 2. However, descriptions of feeding behaviour could well be improved by including more than two components in the mixture (as suggested by Tolkamp and Kyriazakis, 1999) or by the use of other distributions (as suggested by Yeates et al., 2001).

We applied our models to data from cows and interpreted our results in the light of biological expectations based on the satiety principle (as discussed in Tolkamp et al., 1998). Since satiety is thought to be relevant for the feeding behaviour of all species (Simpson, 1990), our conclusions could equally well apply to other species and types of behaviour. An example is male sexual behaviour, for instance as observed in deer mice (Dewsbury, 1983) and rabbits (Gonzalez-Mariscal et al., 1997). Bouts of sexual behaviour are ended when animals are ‘sexually satiated’ while subsequently the probability of another bout gradually increases (Dewsbury, 1983; Gonzalez-Mariscal et al., 1997). Also the breathing patterns of intermittent breathers such as right whales (Winn et al., 1995) and snapping turtles (Frische et al., 2000) occur in bouts, with the probability of a bout occurring increasing with time since the last bout

ended. Therefore it seems likely that our conclusions with respect to the importance of ‘memory’ for the statistical modelling of short-term behaviour are relevant for other types of behaviour that are affected by ‘satiety-like’ phenomena as well.

## 5 CONCLUSIONS

We observe that the (lack of) fit of the three models to the cow feeding data allows us to draw some strong conclusions about the most appropriate type of model for the analysis of short-term feeding behaviour. The fit of the models improved as more allowance was made for effects of the past on future behaviour. We believe this finding is strongly related to the effects of satiation and of (diminishing) satiety on the likelihood of animals finishing and starting a feeding bout, respectively. The hidden Markov model is a ‘current state only’ model based on a ‘lack of memory’ principle and cannot, therefore, allow for satiety effects. This model gave by far the worst description of the observations. Unfortunately, many analyses of short-term feeding (and other types of) behaviour of the last few decades have been based on the Markov principle. Our analyses call for a re-evaluation of the suitability of such models to describe behaviours that are affected by satiety-like principles. The increasing probability of animals initiating a bout played no part in the formulation of the latent Gaussian model, nevertheless the shape of the frequency distributions predicted by this model are consistent with such an increase in probability. However, when a latent structure, based on an explicit description of the frequency distribution of durations between feeding events, was imposed using the semi-Markov model, the best predictions were obtained. This allows the likelihood of animals initiating a bout to change with time and also allows a meal criterion to be determined. We used a mixture of two log-normal distributions to describe the frequency distribution of durations between feeding events; enhancements could be made by considering the use of additional states (e.g. drinking) or by using other distributions that are biologically more appealing. Since satiety is thought to play a role in the short-term feeding behaviour of most species and in several other types of behaviour, our approach may have well much wider applications. For such behaviours, the capacity of statistical models to incorporate effects of long-term ‘memory’ on future probabilities of behavioural events seems indispensable.

**Acknowledgements:** We thank SEERAD for funding this work; DJA also acknowledges his BioSS/ICMS PhD studentship.

## References

Allcroft, D. J. (2001). *Statistical models for short-term animal behaviour*. Ph. D. thesis, University of Edinburgh.

- Allcroft, D. J. and C. A. Glasbey (2002a). A simulation-based evaluation of animal behaviour models. *Statistical Modelling: An International Journal*. (submitted).
- Allcroft, D. J. and C. A. Glasbey (2002b). A spectral estimator of ARMA parameters from thresholded data. *Statistics and Computing*. (in press).
- Box, G. E. P., G. M. Jenkins, and G. C. Reinsel (1994). *Time Series Analysis: Forecasting and Control* (third ed.). Englewood Cliffs, New Jersey: Prentice Hall.
- Collier, G., D. F. Johnson, and C. Mitchell (1999). The relation between meal size and the time between meals: effects of cage complexity and food cost. *Physiology and Behavior* 67, 339–346.
- Dado, R. G. and M. S. Allen (1993). Continuous computer acquisition of feed and water intakes, chewing, reticular motility and ruminal pH of cattle. *Journal of Dairy Science* 76, 1589–1600.
- Dewsbury, D. A. (1983). Recovery from sexual satiety in deer mice (*Peromyscus maniculatis bairdi*). *Journal of Comparative Physiology* 97, 34–42.
- Friggens, N. C., B. L. Nielsen, I. Kyriazakis, B. J. Tolcamp, and G. C. Emmans (1998). Effects of food composition and stage of lactation on the short-term feeding behavior of dairy cows. *Journal of Dairy Science* 81, 3268–3277.
- Frische, S., A. Fago, and J. Altimiras (2000). Respiratory responses to short term hypoxia in the snapping turtle, *Chelydra serpentina*. *Comparative Biochemistry and Physiology A – Molecular and Integrative Physiology* 126, 223–231.
- Gonzalez-Mariscal, G., M. E. Albonetti, E. Cuamatzi, and C. Beyer (1997). Transitory inhibition of scent marking by copulation in male and female rabbits. *Animal Behaviour* 53, 323–333.
- Haccou, P. and E. Meelis (1994). *Statistical Analysis of Behavioural Data: An Approach Based on Time-structured Models*. Oxford: Oxford University Press.
- Kyriazakis, I. and G. C. Emmans (1991). Diet selection in pigs – Dietary choices made by growing pigs following a period of underfeeding with protein. *Animal Production* 52, 337–346.
- Langton, S. D., D. Collett, and R. M. Sibly (1995). Splitting behaviour into bouts; a maximum likelihood approach. *Behaviour* 132, 781–799.
- MacDonald, I. L. and D. Raubenheimer (1995). Hidden Markov models and animal behaviour. *Biometrical Journal* 37, 701–711.
- MacDonald, I. L. and W. Zucchini (1997). *Hidden Markov and Other Models for Discrete-valued Time Series*. London: Chapman & Hall.
- Mayes, E. and P. Duncan (1986). Temporal patterns of feeding behaviour in free-ranging horses. *Behaviour* 96, 105–129.

- Metz, H. (1974). Stochastic models for the temporal fine structure of behaviour sequences. In D. McFarland (Ed.), *Motivational Control Systems Analysis*, pp. 5–86. London: Academic Press.
- Metz, J. H. M. (1975). *Time Patterns of Feeding and Rumination in Domestic Cattle*. Wageningen: Communications Agricultural University 75-12.
- Morgan, C. A., G. C. Emmans, B. J. Tolcamp, and I. Kyriazakis (2000). Analysis of the feeding behavior of pigs using different models. *Physiology and Behavior* 63, 395–403.
- Rook, A. J. and C. A. Huckle (1997). Activity bout criteria for grazing cows. *Applied Animal Behaviour Science* 54, 89–96.
- Sibly, R. M., H. M. R. Nott, and D. J. Fletcher (1990). Splitting behaviour into bouts. *Animal Behaviour* 39, 63–69.
- Simpson, S. J. (1990). The pattern of feeding. In R. F. Chapman and A. Joern (Eds.), *Biology of Grasshoppers*. New York: John Wiley.
- Simpson, S. J., L. Barton Brown, and A. C. M. van Gerwen (1989). The patterning of compensatory sugar feeding in the Australian blowfly. *Physiological Entomology* 14, 91–105.
- Slater, P. J. B. and N. P. Lester (1982). Minimising errors in splitting behaviour into bouts. *Behaviour* 79, 153–161.
- Stamer, E., W. Junge, and E. Kalm (1997). Temporal pattern of feeding of dairy cows kept in groups. *Archiv fuer Tierzuechtung* 40, 195–214. In German.
- Tolcamp, B. J., D. J. Allcroft, E. J. Austin, B. L. Nielsen, and I. Kyriazakis (1998). Satiety splits feeding behaviour into bouts. *Journal of Theoretical Biology* 194, 235–250.
- Tolcamp, B. J., N. C. Friggens, G. C. Emmans, I. Kyriazakis, and J. D. Oldham (2002). Meal patterns of dairy cows consuming diets with a high or a low ratio of concentrate to grass silage. *Animal Science* 74, 369–382.
- Tolcamp, B. J. and I. Kyriazakis (1997). Measuring diet selection in dairy cows: effect of training on selection of dietary protein level. *Animal Science* 64, 197–207.
- Tolcamp, B. J. and I. Kyriazakis (1999). To split behaviour into bouts, log-transform the intervals. *Animal Behaviour* 57, 807–817.
- Tolcamp, B. J., D. P. N. Schweitzer, and I. Kyriazakis (2000). The biologically relevant unit for the short-term feeding behavior of dairy cows. *Journal of Dairy Science* 83, 2057–2068.
- Winn, H. E., J. D. Goosyear, R. D. Kenney, and R. O. Petricig (1995). Dive patterns of tagged right whales in the Great South Channel. *Continental Shelf Research* 15, 593–611.

Yeates, M. P., B. J. Tolkamp, D. J. Allcroft, and I. Kyriazakis (2001). The use of mixture distribution models to determine bout criteria for analysis of animal behaviour. *Journal of Theoretical Biology* 213, 413–425.

**Figure captions:**

Table 1: Background information regarding the data set: mean and variation in individual animal performance during the observation period.

Table 2: Summary of parameter estimates for the semi-Markov models fit to the eight cows.

Table 3: Characteristics of meals as predicted by two of the models.

Figure 1: Two days of feeding data for one of the cows. Raised values of the signal denote periods of feeding.

Figure 2: Diagrammatic representation of the possible model structure.

Figure 3: Realisation of a hidden Markov model. The underlying state (**M**eal or **B**etween-meal) follows a Markov chain. The probability of the observed behaviour being feeding (**1**) or non-feeding (**0**) are dependent on the current underlying state, i.e. **B** or **M**.

Figure 4: Simulations from a latent Gaussian model, fit to one of the cows. The threshold level is chosen to match the observed rate of feeding and the resulting feeding events are shown above.

Figure 5: Distribution of log-transformed non-feeding durations, with fitted mixture of two log-normal distributions.

Figure 6: Marginal distributions for durations of (a) feeding events, (b) non-feeding events. (+) data; (—) hidden Markov model; (· — ·) latent Gaussian model; (···) semi-Markov model.

**Short title:** Models for animal feeding behaviour