

A simulation study for the analysis of uncertain binary responses: Application to first insemination success in beef cattle

Robyn L. SAPP^{a*}, Matthew L. SPANGLER^a, Romdhane REKAYA^{a,b**},
J. Keith BERTRAND^a

^a Animal and Dairy Science Department, The University of Georgia,
Athens, Georgia 30602, USA

^b Department of Statistics, The University of Georgia,
Athens, Georgia 30602, USA

(Received 5 July 2004; accepted 6 August 2005)

Abstract – A simulation was carried out to investigate the methods of analyzing uncertain binary responses for success or failure at first insemination. A linear mixed model that included, herd, year, and month of mating as fixed effects; and unrelated service sire, sire and residual as random effects was used to generate binary data. Binary responses were assigned using the difference between days to calving and average gestation length. Females deviating from average gestation length lead to uncertain binary responses. Thus, the methods investigated were the following: (1) a threshold model fitted to certain (no uncertainty) binary data (M1); (2) a threshold model fitted to uncertain binary data ignoring uncertainty (M2); and (3) analysis of uncertain binary data, accounting for uncertainty from day 16 to 26 (M3) or from day 14 to 28 (M4) after introduction of the bull, using a threshold model with fuzzy logic classification. There was virtually no difference between point estimates obtained from M1, M3, and M4 with true values. When uncertain binary data were analyzed ignoring uncertainty (M2), sire variance and heritability were underestimated by 22 and 24%, respectively. Thus, for noisy binary data, a threshold model contemplating uncertainty is needed to avoid bias when estimating genetic parameters.

binary data / fertility / fuzzy logic / simulation / threshold model

1. INTRODUCTION

Until recently, limited resources had been allocated to the research of fertility measures in beef cattle. This was, in part, due to the limited availability of data and computational requirements for traits of this nature. Two sources

* The first and second author contributed equally to this manuscript.

** Corresponding author: rrekaya@uga.edu

of information, natural service (NS) and artificial insemination (AI) mating data are generally associated with fertility traits. Few studies, however, have looked at the possibility of combining these two sources of information for the purpose of genetic evaluation. Furthermore, these two sources of information have not previously been combined due to the absence of a trait that could be used to describe fertility in both NS and AI mating data. Donoghue *et al.* [5] suggested a binary trait, calving to first insemination (CFI), which evaluated the probability that a calving event occurred as the result of conception at first insemination. For NS matings, success at first insemination was defined as conceiving during the first 21 days of the breeding season, which corresponded to the first estrous cycle of the female.

A binary trait such as CFI, for genetic evaluation of combined NS and AI mating data, as opposed to a continuous trait such as days to calving (DC), is dictated by the nature of the recording scheme currently used in the beef cattle industry; DC is defined as the time elapsed between the introduction of the bull and the subsequent calving date. In an NS mating situation, the date at which the female is exposed is recorded but the conception date is unknown. In contrast, AI data provides an accurate insemination date but the first exposure is uncertain. Yet, Donoghue *et al.* [5] proposed that NS and AI data could be combined in one analysis for the trait of CFI due to the high genetic (0.82) correlation between NS and AI mating data for CFI. Although this result is encouraging, it does not alleviate the uncertainty of CFI in the NS data.

With AI data, the exact day of insemination is known whereas in NS data, the number of days to insemination (NDI) resulting in conception is unknown. In order to derive NDI (the time interval between the introduction of the bull and insemination resulting in conception), knowledge about DC and the length of the estrous cycle is used. Since the gestation length (GL) in NS mating data is unknown, an average GL, obtained from AI mating data is deviated from DC to give the approximate NDI. This procedure lends itself to the possibility of uncertainty when females deviate in GL. For example, if a female has an NDI of 20 days, but has a longer than average GL, she would have been classified as failing to conceive at first insemination when the opposite (successful conception at first insemination) was true; thus leading to uncertain binary responses of success or failure at first insemination.

The current study was motivated by the belief that uncertainty is present in NS mating data for a trait such as CFI (at an unknown rate) and to date there have been no studies addressing this particular issue. One proposed way to account for uncertainty in success or failure at first insemination is *via* a threshold model using fuzzy logic classification, where fuzzy logic classification uses

imprecise propositions based on fuzzy set theory to assign partial membership of a set [2]. Therefore, the objective of the current study was to evaluate different methods of analyzing uncertain binary responses with application to success or failure at first insemination in beef cattle.

2. MATERIALS AND METHODS

2.1. Statistical analysis and computations

Threshold models are becoming a standard tool for the analysis of discrete data in the field of animal breeding and genetics. Extensive literature on its theoretical basis, implementation and application has been generated in the last twenty years [6, 7, 12]. More recently, Rekaya *et al.* [10] have proposed a method for analyzing binary data subject to misclassification using a threshold model. In the present study, an extension of such a method, based on fuzzy logic classification, is presented.

2.1.1. Threshold model for analysis of uncertain binary responses

Let $\mathbf{m} = (m_1, m_2, \dots, m_n)'$ be a sample of uncertain independent binary responses of non-observed real data $\mathbf{y} = (y_1, y_2, \dots, y_n)'$, where each y_i is Bernoulli with success probability p_i that is expressed as a function of some systematic and random effects. Miscoding occurs if some y_i is switched, *e.g.*, $y_i = 0$ becomes $m_i = 1$ (*i.e.*, a zero coded as one). Following notation by Rekaya *et al.* [10], let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)'$, where α_i is an indicator variable for observation i such that $\alpha_i = 1$ if y_i is switched, $\alpha_i = 0$ otherwise. Suppose each α_i is a Bernoulli with success probability π_t (probability of uncertainty) at time t such that $p(\alpha_i|\pi_t) = \pi_t^{\alpha_i}(1 - \pi_t)^{(1-\alpha_i)}$, where π_t is specific for each observation and will be explained further in Section 2.1.2.

Assuming independence between $\boldsymbol{\alpha}$ and \mathbf{y} , their joint distribution, given $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$, is

$$p(\boldsymbol{\alpha}, \mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\pi}) = \prod_{i=1}^n p(\alpha_i|\pi_t)p(y_i|\boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\theta}$ is the vector of systematic and random effects and $\boldsymbol{\pi} = (\pi_{t_1}, \pi_{t_2}, \dots, \pi_{t_n})'$. Thus,

$$p(\boldsymbol{\alpha}, \mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\pi}) = \prod_{i=1}^n \pi_{t_i}^{\alpha_i}(1 - \pi_{t_i})^{(1-\alpha_i)} [p_i(\boldsymbol{\theta})]^{y_i} [1 - p_i(\boldsymbol{\theta})]^{(1-y_i)}, \tag{2}$$

where $p_i(\boldsymbol{\theta}) = \Phi_i(\boldsymbol{\theta})$, and $\Phi_i(\cdot)$ is the standard normal cumulative distribution function. Furthermore, the following relationship between y_i and m_i , given α_i , could be established as

$$y_i = (1 - \alpha_i)m_i + \alpha_i(1 - m_i) \quad (3)$$

such that $y_i = m_i$ when there is no miscoding of the binary response.

Using the relationships in equations (2) and (3), the joint probability of $\boldsymbol{\alpha}$ and \mathbf{m} , given $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$, is

$$p(\boldsymbol{\alpha}, \mathbf{m} | \boldsymbol{\theta}, \boldsymbol{\pi}) = \prod_{i=1}^n \pi_{t_i}^{\alpha_i} (1 - \pi_{t_i})^{(1-\alpha_i)} [p_i(\boldsymbol{\theta})]^{(1-\alpha_i)m_i + \alpha_i(1-m_i)} [1 - p_i(\boldsymbol{\theta})]^{1 - [(1-\alpha_i)m_i + \alpha_i(1-m_i)]}. \quad (4)$$

Finally, prior distributions for $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$ would complete the Bayesian formulation. However, in some situations, $\boldsymbol{\pi}$ is known or could be inferred from external information. In the next section, fuzzy logic classification was used to determine the vector $\boldsymbol{\pi}$.

2.1.2. Application to first insemination success in beef cattle

In beef cattle when natural mating is used, there is no direct recording of conception within the first 21 days of the breeding season. The first 21 days of the breeding season correspond to the length of the estrus cycle in the cattle. Therefore, we are interested in discriminating between cows that conceived in the first 21 days after introduction of the bull — conception occurring during the first estrus cycle — and those that did not conceive in the first 21 days. The only information available from NS matings is DC, which is computed as the time elapsed between the introduction of the bull and the subsequent calving date. Success or failure at first insemination (FIS) is based on the difference between DC and an average GL, where the average GL differed by sex of the calf. Thus, if the difference between DC and average GL is less than or equal to 21 days then FIS = 1, otherwise FIS = 0; in other words, a successful FIS event would be defined as conception occurring during the first 21 days of the breeding season.

Given the variation in GL between cows, and potentially of the estrus cycle, it is possible that some cows have uncertain FIS. Furthermore, the variation in GL, estrus cycle, and the difference between DC and average GL could be used to assess the probability of miscoding for every FIS record. In this simulation study, given the limited information on the variation of the estrus cycle

and for the sake of simplicity, only variability in GL was used to assess the uncertainty of FIS. However, for real data sets and depending on the available information, sources of variation other than GL can be used within the framework of the presented method in a straightforward manner.

In the current study, fuzzy logic classification, based on the binary response of FIS and the difference between DC and average GL, was used to calculate the probability of miscoding at time t_i . The deviation of average GL from DC gives the approximate NDI (the time interval between the introduction of the bull and insemination resulting in conception). It was then assumed that if the absolute value of the deviation of NDI and 21 was smaller (greater) than a number of days (based on practical and biological knowledge), then the observed FIS response was certain; in other words, if only one standard deviation (5 days) of GL is used to assess uncertainty, cows having an absolute difference between DC and average GL of less than 16 days or greater than 26 days, then there was no uncertainty about the observed FIS response. Otherwise, the binary record was considered as uncertain.

To assess the uncertainty about FIS, a fuzzy logic approach was used. There is no analytical method to determine a fuzzy function. However, such functions have to be derived based on practical and biological knowledge about the data generating process and good heuristics. Given the subjective nature of fuzzy classification, three fuzzy functions were used to assess their validity and the robustness of the method to fuzzy function choice. The following three fuzzy logic functions were used to compute the miscoding at time t_i (see Fig. 1):

$$\pi_{t_i} = \begin{cases} 0, & 16 < t_i \\ 0.5 - (21 - t_i)/10, & 16 \leq t_i \leq 21 \\ 0.5 + (21 - t_i)/10, & 21 < t_i \leq 26 \\ 0, & t_i > 26 \end{cases} \quad (5a)$$

$$\pi_{t_i} = \begin{cases} 0, & 16 < t_i \\ 0.5 - (21 - t_i)/16, & 16 \leq t_i \leq 21 \\ 0.5 + (21 - t_i)/16, & 21 < t_i \leq 26 \\ 0, & t_i > 26 \end{cases} \quad (5b)$$

and

$$\pi_{t_i} = \begin{cases} 0, & 16 < t_i \\ -0.02t_i^2 + 0.84t_i - 8.32, & 16 \leq t_i \leq 26 \\ 0, & t_i > 26 \end{cases} \quad (5c)$$

The interval in equations (5a–c) was based on one standard deviation of five days for GL obtained from AI data. Thus, the interval of uncertain FIS responses ranged from day 16 to 26 after introduction of the bull. In order

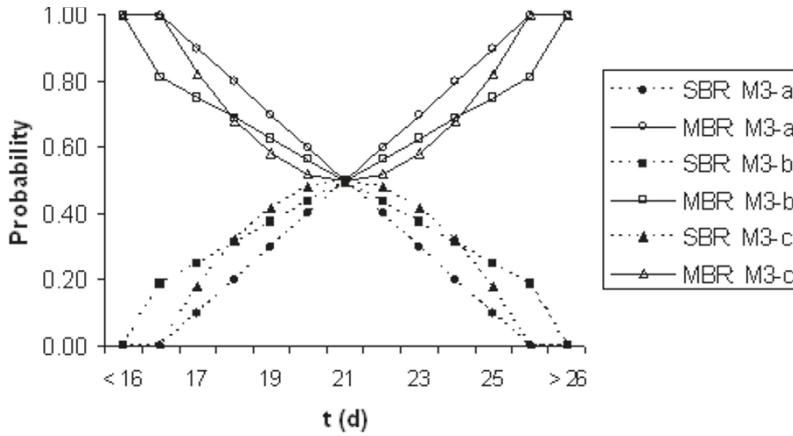


Figure 1. Probability of the observed binary response, given the approximate number of days to insemination (NDI), was maintained (MBR) or switched (SBR) from day 16 to 26 after introduction of the bull using fuzzy logic classification.

to evaluate the robustness of the procedure, the interval of uncertain binary responses was increased from 10 to 14 days (nearly 1.5 standard deviation units) and the fuzzy functions were adjusted accordingly (see Fig. 2):

$$\pi_{t_i} = \begin{cases} 0, & 14 < t_i \\ 0.5 - (21 - t_i)/14, & 14 \leq t_i \leq 21 \\ 0.5 + (21 - t_i)/14, & 21 < t_i \leq 28 \\ 0, & t_i > 28 \end{cases}, \tag{6a}$$

$$\pi_{t_i} = \begin{cases} 0, & 14 < t_i \\ 0.5 - (21 - t_i)/21, & 14 \leq t_i \leq 21 \\ 0.5 + (21 - t_i)/21, & 21 < t_i \leq 28 \\ 0, & t_i > 28 \end{cases}, \tag{6b}$$

and

$$\pi_{t_i} = \begin{cases} 0, & 14 < t_i \\ -0.01t_i^2 + 0.42t_i - 3.92, & 14 \leq t_i \leq 28 \\ 0, & t_i > 28 \end{cases}. \tag{6c}$$

Although the interval of uncertain FIS could be increased to more than 1.5 standard deviation units (cows with very short or long GL), it is biologically reasonable to assume that the number of cows with very short (long) GL is small. Furthermore, increasing the interval of uncertain NDI days to more than 1.5 standard deviation units will unnecessarily increase the number of potential uncertain responses and may not be the best approach.

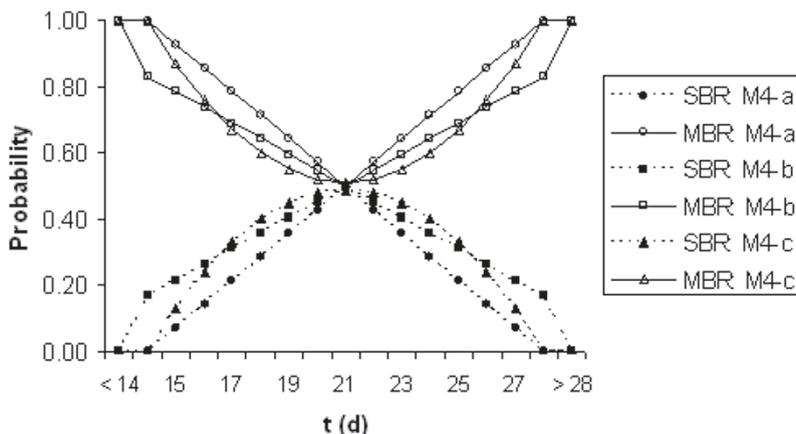


Figure 2. Probability of the observed binary response, given the approximate number of days to insemination (NDI), was maintained (MBR) or switched (SBR) from day 14 to 28 after introduction of the bull using fuzzy logic classification.

A mixed linear model was used for the analysis of the underlying liability of FIS. In matrix notation the model could be written as:

$$\lambda = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_s\mathbf{s} + \mathbf{Z}_u\mathbf{u} + \mathbf{e},$$

where λ is a vector of unobserved liabilities, $\boldsymbol{\beta}$ is the vector of fixed effects, \mathbf{s} is the vector of unrelated random service sire effects, \mathbf{u} is the vector of sire effects, and \mathbf{e} is the vector of residual effects. Furthermore, \mathbf{X} , \mathbf{Z}_s , and \mathbf{Z}_u are the corresponding incidence matrices with the appropriate dimensions.

Using the same notation as in the previous section, let \mathbf{m} be defined as the vector of observed uncertain FIS responses and \mathbf{y} be defined as the vector of unobserved true FIS responses. The joint probability of $\boldsymbol{\alpha}$ and \mathbf{m} , given $\boldsymbol{\theta} = (\boldsymbol{\beta}', \mathbf{s}', \mathbf{u}')'$ and $\boldsymbol{\pi}$, is identical to equation (4) such that

$$p(\boldsymbol{\alpha}, \mathbf{m} | \boldsymbol{\theta}, \boldsymbol{\pi}) = \prod_{i=1}^n \pi_i^{\alpha_i} (1 - \pi_i)^{(1-\alpha_i)} [p_i(\boldsymbol{\theta})]^{(1-\alpha_i)m_i + \alpha_i(1-m_i)} [1 - p_i(\boldsymbol{\theta})]^{1 - [(1-\alpha_i)m_i + \alpha_i(1-m_i)]}, \quad (7)$$

where $p_i(\boldsymbol{\theta}) = \Phi_i(\mathbf{x}'_i\boldsymbol{\beta} + \mathbf{z}'_{si}\mathbf{s} + \mathbf{z}'_{ui}\mathbf{u})$ is the probability of success or failure at first insemination for record i and \mathbf{x}'_i , \mathbf{z}'_{si} , and \mathbf{z}'_{ui} are the known row vectors relating the fixed, service sire and additive sire effects to the probability of first insemination success, respectively.

To ensure proper posterior distribution, the following priors were assumed for the parameters in the model:

$$p(\boldsymbol{\beta}) \sim N(\mathbf{0}, \mathbf{10}^6), \quad (8)$$

$$p(\mathbf{s}|\sigma_s^2) \sim N(\mathbf{0}, \mathbf{I}\sigma_s^2), \quad (9)$$

and

$$p(\mathbf{u}|\mathbf{A}, \sigma_u^2) \sim N(\mathbf{0}, \mathbf{A}\sigma_u^2), \quad (10)$$

where \mathbf{I} is the identity matrix and \mathbf{A} is a known matrix of relationships between sires. Uniform bounded priors, $U[0, 1]$, were assumed for σ_s^2 and σ_u^2 .

The joint posterior density was obtained by the product of densities in equations (7) through (10) leading to

$$p(\boldsymbol{\theta}, \sigma_s^2, \sigma_u^2, \boldsymbol{\alpha}|\mathbf{m}, \boldsymbol{\pi}) \propto \prod_{i=1}^n \pi_{t_i}^{\alpha_i} (1 - \pi_{t_i})^{(1-\alpha_i)} [p_i(\boldsymbol{\theta})]^{(1-\alpha_i)m_i + \alpha_i(1-m_i)} [1 - p_i(\boldsymbol{\theta})]^{1 - [(1-\alpha_i)m_i + \alpha_i(1-m_i)]} p(\boldsymbol{\beta}) p(\mathbf{u}|\mathbf{A}, \sigma_u^2) p(\mathbf{s}|\sigma_s^2)$$

defined only within the boundaries of the bounded priors for σ_s^2 and σ_u^2 .

Following Rekaya *et al.* [10], the conditional posterior distribution of each α_i was given by

$$p(\alpha_i = 1|\boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\alpha}_{-i}, \sigma_u^2, \sigma_s^2, \mathbf{m}) = \frac{\pi_{t_i} [p_i(\boldsymbol{\theta})]^{(1-m_i)} [1 - p_i(\boldsymbol{\theta})]^{m_i}}{K} \quad (11)$$

and

$$p(\alpha_i = 0|\boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\alpha}_{-i}, \sigma_u^2, \sigma_s^2, \mathbf{m}) = \frac{(1 - \pi_{t_i}) [p_i(\boldsymbol{\theta})]^{m_i} [1 - p_i(\boldsymbol{\theta})]^{(1-m_i)}}{K}, \quad (12)$$

where $K = \pi_{t_i} [p_i(\boldsymbol{\theta})]^{(1-m_i)} [1 - p_i(\boldsymbol{\theta})]^{m_i} + (1 - \pi_{t_i}) [p_i(\boldsymbol{\theta})]^{m_i} [1 - p_i(\boldsymbol{\theta})]^{(1-m_i)}$. Thus, each α_i ($i = 1, 2, \dots, n$) was sampled from a discrete distribution, with probabilities as in equations (11) and (12). Once all α_i ($i = 1, 2, \dots, n$) had been sampled, the true data, \mathbf{y} , was generated using the relationship in equation (3). The joint posterior density of $\boldsymbol{\theta}$, σ_s^2 , and σ_u^2 , given the true data, was obtained as follows:

$$p(\boldsymbol{\theta}, \sigma_s^2, \sigma_u^2|\mathbf{y}) \propto \prod_{i=1}^n p_i(\boldsymbol{\theta})^{y_i} [1 - p_i(\boldsymbol{\theta})]^{(1-y_i)} p(\boldsymbol{\beta}) p(\mathbf{u}|\mathbf{A}, \sigma_u^2) p(\mathbf{s}|\sigma_s^2). \quad (13)$$

Table I. Summary of the parameters used in the general and the field data based simulation.

Item	General	Field data
Herds	100	275
Years	13	13
Months of calving	12	12
Service sires	300	1 504
Daughter records	10 000	24 961
Sires with daughter records	460	1 667
Sires in pedigree file	500	2 397

The joint posterior density in equation (13) was augmented with the liabilities for all observations in the data set. All conditional posterior distributions of model parameters were in closed form as described by Albert and Chib [1] and Sorensen *et al.* [12]. These distributions were normal for the location parameters, truncated normal for each of the liabilities and scaled-inverted chi-square distributions for the dispersion parameters. Liabilities were sampled from their truncated normal distribution using the inverse cumulative distribution function technique [4].

For each round of Gibbs sampling, α_i was sampled, taking the value of one or zero. At completion of the sampling process, the probability that observation i was uncertain could be estimated as the ratio between the number of times that α_i was one and the total number of samples.

2.2. General simulation

A simulation, using a sire model, was carried out to investigate two methods of analyzing uncertain binary responses for FIS. Five over-lapping generations were simulated. The base population included forty unrelated sires and subsequent generations consisted of 115 sires each. Thus, a total of 500 sires were generated. The data set consisted of 10 000 daughter records from bulls in generations 2 through 5. Fixed and random effects were randomly assigned to each record. In all cases, no extreme case (ECP) classes were allowed. A more detailed explanation of ECP classes and the effects on variance component estimation can be found in Moreno *et al.* [8]. A summary description of the simulated data and pedigree is presented in Table I.

A linear mixed model at the liability scale which included herd, year, and month of mating as fixed effects; and unrelated service sire, sire and residual as random effects were used to generate the binary responses. Herd and

year/month of mating effects were drawn from uniform distributions, $U[0.4, 0.9]$ and $U[-0.4, -0.3]$, respectively. Service sire effects were drawn from a normal distribution, $N(\mathbf{0}, \mathbf{I}\sigma_s^2)$, where \mathbf{I} is the identity matrix and $\sigma_s^2 = 0.03$. Transmitting abilities of the base population sires were sampled from a normal distribution, $N(\mathbf{0}, \mathbf{I}\sigma_u^2)$, where \mathbf{I} is the identity matrix and $\sigma_u^2 = 0.02$. The remaining sire effects were sampled from a normal distribution with the mean equal to one half of the grandsire's transmitting ability and variance equal to $(3/4)\sigma_u^2$. The residual variance was sampled from a normal distribution, $N(\mathbf{0}, \mathbf{I}\sigma_e^2)$, where \mathbf{I} is the identity matrix and $\sigma_e^2 = 1.0$.

The liability for each daughter record was calculated as the sum of all effects included in the model. Given the liability, the cycle in which the daughter conceived was assigned. A daughter was randomly assigned to one of five cycles, where cycle 1 is the first 21 days after introduction of the bull (*i.e.*, the first estrous cycle occurring after the bull was introduced), cycle 2, 3, and 4 were the second, third and fourth 21 days, respectively, after the introduction of the bull, and cycle 5 was non-calvers (failed to calve). The NDI was randomly sampled for cows conceiving in cycles 1 to 4 using the integer value in the following formula:

$$\text{NDI} = [21(c - 1) + 21u] + 1 \quad (14)$$

where c is the cycle and u was drawn from a uniform distribution, $U[0, 1]$.

The NDI obtained in equation (14) was then used to generate two data sets. The first data set consisted of binary responses (D1) for FIS assigned using simulated NDI, where the binary responses were set to one if NDI was less than or equal to 21 days, and zero otherwise (including non-calvers). The second data set (D2) consisted of DC values. The DC values were obtained by generating a GL from a normal distribution with mean (standard deviation) of 280 (5.1) days for bull calves and 279 (5.2) days for heifer calves and adding the simulated NDI from equation (14). Sex of calf was randomly assigned and non-calvers (cycle 5) were assigned a subjective DC value equal to 980 days.

To mimic the field data, DC records in D2, assuming that NDI and GL were unknown, were used to generate the binary trait of FIS. For that purpose, an average GL specific to each sex (280 and 279 days for male and female calves, respectively) was subtracted from the DC in D2. If the difference was less than or equal to 21 days, then $\text{FIS} = 1$, otherwise $\text{FIS} = 0$ (including non-calvers). This third data set was labeled D3.

Three methods of accounting for uncertainty in success or failure at first insemination were investigated: (1) analysis of D1 (certain data – no uncertainty present) using a threshold model (M1); (2) analysis of D3 (uncertain data), without consideration of uncertainty, using a threshold model (M2); and (3)

analysis of D3 (uncertain data), using a threshold model with fuzzy logic classification (M3 and M4). Three different fuzzy logic functions were used to account for uncertain FIS responses. For M3, the use of fuzzy function in equations (5a), (5b), and (5c) to account for uncertainty will be denoted as M3-a, M3-b, and M3-c, respectively. Similarly, the use of equations (6a), (6b), and (6c), in M4, will be denoted as M4-a, M4-b, and M4-c, respectively. Ten replicates were simulated for analysis using M1, M2, M3a-c, and M4a-c.

2.3. Field data based simulation

In order to evaluate the validity of the proposed method with field data, a real data based simulation, using a sire model, was carried out. The structure of a true available data set after stringent editing was used for the simulation. Editing consisted in removing herds and service sires with less than five records as well as daughters with unknown sires. The resulting data set consisted of 24 961 daughter records from 1 667 sires. A description of the field data structure and pedigree is presented in Table I.

A linear mixed model which included herd, year, and month of mating as fixed effects; and unrelated service sire, sire and residual as random effects were used to generate the binary responses. Herd, year, and month of mating effects were drawn from uniform distributions, $U[0.10, 0.60]$, $U[-0.35, -0.25]$, and $U[-0.40, -0.30]$, respectively. Service sire, sire, and residual effects were sampled in the same manner as the original simulation in Section 2.2.

A daughter was randomly assigned to one of five cycles using the sum of all effects included in the model, where the five cycles were defined as in Section 2.2. The NDI was randomly sampled using equation (14) and three data sets (D1, D2, and D3) were generated as in Section 2.2. The simulated data was then analyzed using the three methods (M1, M2, and M3) exactly in the same manner as in Section 2.2. For M3 and given the lack of differences between the three fuzzy functions used in the first simulation, only one fuzzy function (5a) was used to account for uncertainty with the real data based simulation and will be denoted as M3a. Ten replicates were simulated.

2.4. Convergence

Convergence diagnostics were based on the method of Raftery and Lewis [9] as implemented in the Bayesian Output Analysis (BOA) program [11]. The required burn-in period was always less than 2000 iterations for all parameters and all analyses. Thus, a total chain length of 75 000 iterations of the Gibbs

sampler was run with a conservative burn-in of 25 000 iterations. The remaining 50 000 iterations were retained without thinning for post-Gibbs analysis. Point estimates of the mean and standard deviation were obtained from the BOA software [11] for all parameters and analyses. Furthermore, the high posterior density 95% [HPD (95%)] intervals were calculated using the algorithm described by Chen and Shao [3] as implemented in the BOA program [11].

3. RESULTS

3.1. General simulation

3.1.1. Variance components

The posterior mean, standard deviation and the HPD (95%) interval for service sire and sire variance, based on 10 replicates, is presented in Table II. The point estimates of service sire (0.033) and sire (0.018) variance using M1 were, as expected, similar to the true values (0.03 and 0.02, respectively) used in the simulation. For M2, point estimates of the service sire (0.026) and sire variance (0.014) were underestimated compared to the true values of these parameters and the point estimates using M1 (0.033 and 0.018, respectively). Furthermore, the true values were barely included in the HPD (95%) interval of M2 indicating a potential bias in the estimation of genetic parameters. The point estimate for sire variance obtained from M3a-c as well as M4a-c (ranging from 0.018 to 0.020) were virtually the same as M1 (0.018), whereas the service sire estimates were slightly smaller (ranging from 0.028 to 0.029) when compared with M1 (0.033). For both parameters, the point estimates were similar to the true values used in the simulation.

3.1.2. Heritability estimates

The posterior mean, standard deviation and the HPD (95%) interval for heritability, based on 10 replicates, is presented in Table III. The point estimate of heritability obtained using M1 (0.070) was slightly lower than the true value (0.076) used in the simulation. The estimate of heritability obtained from M2 (0.053) was clearly underestimated. The true value of heritability (0.076) was barely included in the HPD (95%) interval similarly to the variance components. Using M3a-c and M4a-c, the point estimates of heritability (ranging from 0.070 to 0.0740) were virtually the same as the true value used in the simulation and the point estimate obtained using M1 (0.070).

Table II. Summary of the posterior mean (PM), standard deviation (PS), and lower (HL) and upper (HU) bounds of the high posterior density 95% interval for variance components obtained from the general simulation^a.

Method ^c	σ_s^2 (0.030) ^b				σ_u^2 (0.020)			
	PM	PS	HL	HU	PM	PS	HL	HU
M1	0.033	0.007	0.019	0.047	0.018	0.005	0.009	0.029
M2	0.026	0.006	0.014	0.039	0.014	0.005	0.005	0.023
M3a	0.028	0.007	0.015	0.041	0.019	0.005	0.011	0.028
M3b	0.029	0.007	0.015	0.042	0.019	0.005	0.011	0.029
M3c	0.028	0.007	0.015	0.042	0.018	0.005	0.010	0.027
M4a	0.028	0.007	0.016	0.042	0.019	0.005	0.011	0.029
M4b	0.029	0.007	0.016	0.043	0.020	0.005	0.011	0.029
M4c	0.029	0.007	0.015	0.043	0.019	0.005	0.010	0.028

^a The values reported are the average of 10 replicates.

^b σ_s^2 = service sire variance; σ_u^2 = sire variance; and the values in parentheses are the true values used in the simulation.

^c M1 = analysis of binary data using a threshold model; M2 = analysis of uncertain binary data, without consideration for uncertainty, using a threshold model; M3 = analysis of uncertain binary data, accounting for uncertainty from day 16 to 26 after introduction of the bull, using a threshold model with one of three fuzzy logic classification functions (functions a, b or c); and M4 = analysis of uncertain binary data, accounting for uncertainty from day 14 to 28 after introduction of the bull, using a threshold model with one of three fuzzy logic classification functions (functions a, b or c).

3.1.3. Pearson correlations

The results of the Pearson correlations for service sire effects and sire breeding values were averaged over the 10 replicates and are presented in Table IV. Pearson correlations between true breeding values and predicted breeding values obtained from M1 and M2 were 0.50 and 0.45, respectively. When the model contemplated uncertainty using fuzzy logic classification (M3a-c and M4a-c), Pearson correlations were 0.46 between true and predicted breeding values. For service sire effects, Pearson correlations between true and estimated service sire effects were 0.61 and 0.58 for M1 and M2, respectively. Pearson correlations were 0.58 between true and estimated service sire effects when the model (M3a-c and M4a-c) contemplated uncertainty using fuzzy logic classification.

Table III. Summary of the posterior mean (PM), standard deviation (PS), and lower (HL) and upper (HU) bounds of the high posterior density 95% interval for heritability obtained from the general simulation^a.

Method ^c	h^2 (0.076) ^b			
	PM	PS	HL	HU
M1	0.070	0.020	0.033	0.109
M2	0.053	0.017	0.021	0.087
M3a	0.073	0.017	0.041	0.107
M3b	0.073	0.018	0.041	0.108
M3c	0.070	0.017	0.038	0.104
M4a	0.073	0.018	0.041	0.108
M4b	0.074	0.018	0.041	0.109
M4c	0.071	0.017	0.039	0.105

^a The values reported are the average of 10 replicates.

^b h^2 = heritability, calculated as $\frac{(4 \cdot \sigma_u^2)}{[1 + \sigma_s^2 + \sigma_u^2]}$; and the value in parentheses is the true value used in the simulation.

^c M1 = analysis of binary data using a threshold model; M2 = analysis of uncertain binary data, without consideration for uncertainty, using a threshold model; M3 = analysis of uncertain binary data, accounting for uncertainty from day 16 to 26 after introduction of the bull, using a threshold model with one of three fuzzy logic classification functions (functions a, b or c); and M4 = analysis of uncertain binary data, accounting for uncertainty from day 14 to 28 after introduction of the bull, using a threshold model with one of three fuzzy logic classification functions (functions a, b or c).

3.2. Field data based simulation

3.2.1. Variance components

The posterior mean, standard deviation and the HPD (95%) interval for service sire and sire variance, based on 10 replicates, is presented in Table IV. The point estimates of service sire (0.033) and sire (0.018) variance using M1 were, as expected, similar to the true values (0.030 and 0.020, respectively) used in the simulation. The point estimates for sire and service sire variance obtained from M3a were 0.017 and 0.030, respectively. The estimates obtained from M3a were quite similar to the corresponding estimates obtained from M1 (0.018 and 0.033, respectively). For both parameters, the point estimates from M1 and M3a were similar to the true values used in the field data based simulation. Similar to the general simulation, point estimates of the service sire (0.023) and sire variance (0.013) obtained from M2 were underestimated compared to the true values of these parameters and the point estimates using M1

Table IV. Summary of the posterior mean (PM), standard deviation (PS), and lower (HL) and upper (HU) bounds of the high posterior density 95% interval for variance components obtained using the field data based simulation^a.

Method ^c	σ_s^2 (0.030) ^b				σ_u^2 (0.020)			
	PM	PS	HL	HU	PM	PS	HL	HU
M1	0.033	0.005	0.023	0.044	0.018	0.004	0.011	0.025
M2	0.023	0.005	0.014	0.033	0.013	0.003	0.007	0.020
M3a	0.030	0.006	0.019	0.041	0.017	0.004	0.009	0.025

^a The values reported are the average of 10 replicates.

^b σ_s^2 = service sire variance; σ_u^2 = sire variance; and the values in parentheses are the true values used in the simulation.

^c M1 = analysis of binary data using a threshold model; M2 = analysis of uncertain binary data, without consideration for uncertainty, using a threshold model; and M3a = analysis of uncertain binary data, accounting for uncertainty from day 16 to 26 after introduction of the bull, using a threshold model with fuzzy logic classification function a.

(0.033 and 0.018, respectively). Furthermore, the true value of service sire and sire variance were barely included in their respective HPD (95%) intervals using M2 indicating a potential bias estimating these parameters.

3.2.2. Heritability estimates

The posterior mean, standard deviation and the HPD (95%) interval for heritability, based on 10 replicates, is presented in Table V. Similar to the general simulation, the point estimate of heritability obtained using M1 (0.069) was slightly lower than the true value (0.076) used in the simulation. The estimate of heritability obtained from M2 (0.051) was clearly underestimated. In fact, the true value of heritability (0.076) was similar to the upper bound of the HPD (95%) interval. Using M3a, the point estimate of heritability (0.065) was slightly smaller than the true value used in the simulation, but similar to the point estimate obtained using M1 (0.069).

3.2.3. Pearson correlations

The results of the Pearson correlations for service sire effects and sire breeding values were averaged over the 10 replicates and are presented in Table VI for the simulation using a field data structure. Pearson correlations between

Table V. Summary of the posterior mean (PM), standard deviation (PS), and lower (HL) and upper (HU) bounds of the high posterior density 95% interval for heritability obtained using the field data based simulation^a.

Method ^c	h^2 (0.076) ^b			
	PM	PS	HL	HU
M1	0.069	0.014	0.042	0.096
M2	0.051	0.013	0.028	0.076
M3a	0.065	0.015	0.036	0.096

^a The values reported are the average of 10 replicates.

^b h^2 = heritability, calculated as $\frac{(4\sigma_u^2)}{[1+\sigma_s^2+\sigma_a^2]}$; and the value in parentheses is the true value used in the simulation.

^c M1 = analysis of binary data using a threshold model; M2 = analysis of uncertain binary data, without consideration for uncertainty, using a threshold model; and M3a = analysis of uncertain binary data, accounting for uncertainty from day 16 to 26 after introduction of the bull, using a threshold model with fuzzy logic classification function a.

true and predicted breeding values obtained from M1, M2, and M3 were 0.40, 0.36, and 0.37, respectively. The Pearson correlations between true and estimated service sire effects were slightly higher than the correlations between breeding values. The correlations between true and estimated service sire effects obtained from M1, M2, and M3 were 0.46, 0.40, and 0.41, respectively.

4. DISCUSSION

4.1. Parameter estimation

The results of this study indicate that a statistical model that contemplates uncertainty of binary responses (M3a-c and M4a-c) resulted in more accurate estimation of variance components and heritability when compared with a standard threshold model (M2). There was virtually no difference between the point estimates obtained from M3a-c and M4a-c with the true values of the parameters (Tabs. II and III). Furthermore, point estimates of genetic parameters using the correct binary data (M1) and the uncertain data analyzed with fuzzy logic classification (M3a-c and M4a-c) were practically the same.

When uncertain binary data were analyzed with a standard threshold model with no consideration for uncertainty (M2), sire variance, service sire variance, and heritability were underestimated by 22, 21, and 24%, respectively, when compared with M1. The service sire variance using M2 was slightly less

Table VI. Summary of the Pearson correlations between true (TR) and estimated additive sire and service sire effects^a.

Effect	TR-M1 ^b	TR-M2	TR-M3a	TR-M3b	TR-M3c	TRM4a	TR-M4b	TR-M4c
General simulation								
Additive	0.498	0.451	0.456	0.459	0.461	0.458	0.462	0.458
Service Sire	0.614	0.577	0.582	0.584	0.582	0.584	0.586	0.578
Field data based simulation								
Additive	0.402	0.364	0.370	–	–	–	–	–
Service Sire	0.462	0.397	0.407	–	–	–	–	–

^a The values reported are the average of 10 replicates.

^bM1 = analysis of binary data using a threshold model; M2 = analysis of uncertain binary data, without consideration for uncertainty, using a threshold model; M3 = analysis of uncertain binary data, accounting for uncertainty from day 16 to 26 after introduction of the bull, using a threshold model with one of three fuzzy logic classification functions (functions a, b or c); and M4 = analysis of uncertain binary data, accounting for uncertainty from day 14 to 28 after introduction of the bull, using a threshold model with one of three fuzzy logic classification functions (functions a, b or c).

affected by the uncertainty of the binary responses. This result was expected because of the large number of records per service sire (on average, 33 records per service sire). In fact, the true value for service sire variance (0.030) was well centered within the HPD (95%) interval for M2. The true values for sire variance (0.020), and heritability (0.076), however, were closer to the upper bound of the corresponding HPD (95%) interval using M2. Based on the normal like shape of the posterior distributions of these parameters, this result indicates that the estimates of sire variance and heritability using model M2 are less likely to be equal to the true values. Thus, point estimates of these parameters will underestimate the true values and could be biased. Conversely, the true value of sire variance, service sire variance, and heritability were well centered within the HPD (95%) interval using M1, M3a-c, and M4a-c; furthermore indicating that a threshold model contemplating uncertain binary responses resulted in a more accurate estimation of genetic parameters.

Similar results were obtained when the simulation was conducted following the structure of an existing field data set. In fact, the statistical model that contemplated uncertainty of binary responses (M3a) resulted in a more accurate estimation of variance components and heritability when compared with a standard threshold model (M2). There was very little difference between the point estimates obtained from M3a with the true values of the parameters (Tabs. IV and V). Furthermore, point estimates of genetic parameters using the correct binary data (M1) and the uncertain data analyzed with fuzzy logic classification (M3a) were practically the same. However, when uncertain binary data were analyzed with a standard threshold model with no consideration for uncertainty (M2), sire variance, service sire variance, and heritability were underestimated by 28, 30, and 26%, respectively, when compared with M1. Sire and service sire variances using M2 were affected by the uncertainty of the binary responses. The underestimation of the genetic parameters was more pronounced in the field data case in part due to the smaller number of records per service sire and sire (on average, 16.6 records per service sire and 15 records per sire) when compared with the general simulation (on average, 33 records per service sire and 21.7 records per sire). Furthermore, the true value for service sire variance (0.030) was barely contained within the HPD (95%) interval for M2. The true value for sire variance (0.020), and heritability (0.076), however, were similar to the upper bound of the corresponding HPD (95%) interval using M2. These results suggest that the estimates of the genetic parameters using model M2 are likely to be biased.

The results of this study suggest that when analyzing binary data with uncertainty, a standard threshold model could lead to biased inferences. The problem

could be avoided, or at least attenuated, by a statistical model that contemplates uncertainty. Moreover, for field data with low heritability, the use of external information is helpful in discriminating between uncertain, which is coded correctly or incorrectly, and miscoded data. In this study, the mean and standard deviation of the GL were used to assess the probability of miscoding of success or failure at first insemination *via* fuzzy logic classification. There did not appear to be large differences between the intervals (14 to 28 days *vs.* 16 to 26 days) or fuzzy logic functions (two linear and one non-linear) used to account for uncertainty of success or failure at first insemination. Furthermore, if extra information is available on potential other sources of uncertainty of FIS, such information should be included in the fuzzy function.

4.2. Pearson correlations

Although there were no observable differences in the Pearson correlations (Tab. VI) between true and estimated effects using M1, M2, and M3 in the current study, we would expect differences to be more noticeable in field data when the numbers of records per service sire or sire is limited. In such a case, a change in the status of one observation could have a major impact on the estimated effects. In NS mating data for small herds, it would not be unreasonable to expect service sires and sires to have fewer numbers of records per sire than those observed in the current study.

Although there were very small differences in the Pearson correlations (Tab. VI) between true and estimated effects using M2 and M3a using a field data structure in the simulation, the correlations were smaller than those obtained from the general simulation. As with the parameter estimation, this could be due to the fact that the number of records per service sire or sire is limited. Thus, a change in the status of one observation could have a more important impact on the estimated effects.

REFERENCES

- [1] Albert J.H., Chib S., Bayesian analysis of binary and polychotomous response data, *J. Am. Statist. Assoc.* 88 (1993) 669–679.
- [2] Chen G., Pham T.T., Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems, CRC Press LLC, Florida, 2001.
- [3] Chen M.-H., Shao Q.-M., Monte Carlo estimation of Bayesian credible and HPD intervals, *J. Comput. Graph. Stat.* 8 (1999) 69–92.
- [4] Devroye L., Non-uniform random variate generation, Springer-Verlag, New York, 1986.

- [5] Donoghue K.A., Rekaya R., Bertrand J.K., Misztal I., Genetic evaluation of calving to first insemination using natural and artificial insemination mating data, *J. Anim. Sci* 82 (2004) 362–367.
- [6] Gianola D., Theory and analysis of threshold characters, *J. Anim. Sci.* 54 (1982) 1079–1095.
- [7] Gianola D., Foulley J.L., Sire evaluation for ordered categorical data with a threshold model, *Genet. Sel. Evol.* 15 (1983) 201–223.
- [8] Moreno C., Sorensen D., García-Cortés L.A., Varona L., Altarriba J., On biased inferences about variance components in the binary threshold model, *Genet. Sel. Evol.* 29 (1997) 145–160.
- [9] Raftery A.E., Lewis S., How many iterations in the Gibbs sampler?, in: Bernardo J.M., Berger J.O., Dawid A.P., Smith A.F.M. (Eds.), *Bayesian Statistics 4*, Oxford University Press, Oxford, 1992, pp. 763–773.
- [10] Rekaya R., Weigel K.A., Gianola D., Threshold model for misclassified binary responses with application to animal breeding, *Biometrics* 57 (2001) 1123–1129.
- [11] Smith B.J., *Bayesian output analysis program (BOA) user's manual*, version 1.1, University of Iowa, Iowa City, 2005.
- [12] Sorensen D.A., Andersen S., Gianola D., Korsgaard I., Bayesian inference in threshold using Gibbs sampling, *Genet. Sel. Evol.* 27 (1995) 229–249.