BAYESIAN AND FREQUENTIST APPROACHES FOR FITTING THE GAMMA-TIME-DEPENDENT MODEL TO DESCRIBE NEUTRAL DETERGENT FIBER DEGRADATION

Hugo Colombarolli BONFÁ[1]
Edenio DETMANN[1]
Fabyano Fonseca e SILVA[1]
Janderson Florêncio FIGUEIRAS[1]

- ABSTRACT: The aim of the study was evaluate and compare the efficiency of Bayesian and frequentist approach to describe the rumen degradation of NDF. Simulated data was composed by four scenarios: regular restriction in the number of incubation times, random loss of incubation times, loss of specific parts of degradation curves, variation in the precision of the incubations procedures. Two real datasets was used, these real data encompassed the evaluation of NDF degradation of a tropical grass (*Brachiaria decumbes*). The model was fitted according their characteristics approach and compared by plots and assessors. The Bayesian and frequentist approach presented reliable estimates of degradation parameters for the majority of the data tested. Therefore, in specific cases with short random records number, the Bayesian approach showed greater bias of the estimates of incubation residue and estimates of degradation rate without a biological coherence of the parameters, compared to frequentist inference. In another words, the Bayesian approach fitted with prior diffuse, presented less flexible. Nevertheless, it is emphasized the importance of the background information before the modeling, mainly for the Bayesian approach, in order to define proper prior distributions. Future thorough studies about the influence of non-informative prior for the parameters are necessary.

- KEYWORDS: Degradation curves; *in situ* method; inference; ruminal degradation.

1 Introduction

The utilization of the frequentist inference, or the classical inference, was almost unanimous among the scientists in the early years of twentieth century. However, with the computational improving, the Bayesian inference reappeared as a viable alternative to statistical modeling and analysis. Bayesian inference was avoided by researchers for a long time because of the highly complex mathematical resolution, which was not considered viable to be made by using simple algebraic algorithms (LESAFFRE and LAWSON, 2012). However, at early 1960's, the Bayesian inference reappeared in a theoretical paper (JEFFREYS, 1961), but just became widely available to be used from 1990's (GELFAND *et al.*, 1990), when complex integration resolutions could be solved by simulation.

One of the most important objectives of the statistics is to fit or build models. When classical frequentist inference is applied on data, population parameters are considered fixed effects; so the information about the parameters is only obtained by sampling. However, when using Bayesian inference, the population parameters are considered random effects and can be described probabilistically. Therefore, for the process of Bayesian estimation, a smaller dataset would be required than do the frequentist inference. In other words, the model fit by Bayesian method presents lower dependence from data size compared to the frequentist approach (BEAUMONT and RANNALA, 2004). Recently, it has got more difficult to perform experiments with a great number of animals because the high labor, high cost and, mainly, the stricter ethical regulations. In a particular way, experiments where a surgical intervention is necessary have been more affected by the ethical committee exigencies. From this, it can be stated that evaluation of the rumen kinetic parameters (*i.e.,* degradation and transit) is one of the most affected evaluations, because the majority of the experiments are performed by using *in situ* methodology. Therefore, more flexible statistical approaches, in terms of number of experimental animals, have been demanded.

When studies on ruminal kinetics are performed in the tropics, the degradation of the neutral detergent fiber (NDF) must be considered one of the most relevant information as NDF is the main source of energy for cattle production (DETMANN *et al.*, 2008) and around 90 to 95% of its utilization occur in the rumen (HUHTANEN *et al.*, 2010). The ruminal degradation pattern of NDF is described by nonlinear models and follows the action-mass law. The rumen kinetics of NDF is typically a time-dependent process as the probability of a fiber particle either escapes to the lower-gut or be degraded by microorganisms varies along the residence time in the rumen. The inclusion of time-dependency of the rumen kinetics in the nonlinear description is most commonly incorporated by using gamma time-dependent models (ELLIS *et al.*, 1994).

However, the adjusting of nonlinear models can be complexed, mainly for rumen kinetics data, because the errors may not follow a normal distribution. Sometimes, the normal distribution is supposed to be asymptotically achieved. However, the actual data number is normally low. On the other hand, the pattern of degradation or transit is described along time, which creates a typical heteroscedasticity or a "funnel" effect (DETMANN *et al.*, 2001) and a dependency between errors. As an alternative to the frequentist approach, the Bayesian methodology does not require the assumption of normality as a necessary condition and the inferences on the parameters are made on their *posterior* distribution. In this case, a model is supposed for each dataset and the parameters of each model are compared based on their *posterior* distributions (ROSSI *et al.*, 2010).

However, it seems necessary to evaluate the model adjustment capacity and to compare the parameter estimates obtained by Bayesian and frequentist approach, in order to clarify the advantages, disadvantages, and limitations of each approach for ruminal degradation modeling. Therefore, the aim of this study was to evaluate and to compare the efficiency of the Bayesian and frequentist approaches to describe the ruminal degradation of neutral detergent fiber by using a gamma time-dependent model.

2 Materials and methods

2.1   Model applied to interpret ruminal degradation of NDF

The basic model applied to interpret the ruminal degradation profile of NDF was based on a Gamma-2 order of time-dependency according to derivations of Van Milgen *et al.* (1991):

$$R(t) = U \times (1 + c \times t) \times \exp(-c \times t) + I, \qquad (1)$$

where $R(t)$ is the residue of NDF at incubation time t (%); U is the potentially degradable fraction of NDF (%); c is the common rate of lag and degradation (h^{-1}); and I is the asymptote reached when $t \rightarrow \infty$, which means the undegradable fraction of NDF (%).

It must be noted that parameters U and I represent the two fractions of the NDF, and their sum must reproduce the total NDF (100%), as this fiber analytical approach does not encompass a soluble fraction (MERTENS, 2005).

2.2   Simulated data

The simulated data used for the model adjusting and evaluation by the two statistical approaches consisted of a main database split into four scenarios and four subsets for each scenario. Each complete scenario consisted of 145 observations simulated according to the model descried in Equation (1). The values of the parameters of the Equation (1) were assumed to be the average values obtained in a previous work carried out by Figueiras (2013), who evaluated the degradation of NDF from a tropical grass in cattle under grazing. The average values were 62.92, 37.08, and 0.059 for U, I, and c, respectively. Each subset was simulated according to a theoretical hour-by-hour incubation design, then varying from 0 to 144 hours of incubation. From the whole subsets ($n = 145$), different scenarios were produced as described below.

Scenario A – Regular restriction in the number of incubation times: the subsets were produced to simulate a decrease in the number of incubation times; however, presenting an equal time interval between incubation points. Therefore, the four subsets represented sampling designs with degradation times equally spaced in 3, 6, 12 and 24 hours (subsets A_1, A_2, A_3, and A_4, respectively). In this way, these subsets were composed by 48, 24, 12 and, 6 incubation points, respectively. To simulate this scenario, the standard deviation of the residual random error among incubation points was assumed to be equal to 1.0.

Scenario B – Random loss of incubation times: the subsets were produced to simulate different degrees of random losses of incubation times. Therefore, the four subsets represented sampling designs with degradation times randomly distributed from 0 to 144 hours of incubation. In this sense, these subsets were composed by 48, 24, 12 and 6 incubation points (subsets B_1, B_2, B_3, and B_4, respectively). To simulate this scenario, the standard deviation of the residual random error among incubation points was assumed to be equal to 1.0.

Scenario C – Loss of specific parts of the degradation curves: the subsets were produced to simulate the loss of incubation points located at the same part of the degradation curve. This scenario was composed by four subsets (C_1, C_2, C_3, and C_4), where each subset consisted of 75% of the total subset record number. Each subset was divided into four parts (four quarters); three of these parts composed each subset. Therefore, each subset have 108 records, where the first subset consist in the main database without the first part (without initial quarter), the second subset consist, of the main database without the second part (without second quarter) and so on. To simulate this scenario, the standard deviation of the residual random error among incubation points was assumed to be equal to 1.0.

Scenario D – Variation in the precision of the incubation procedures: the subsets were produced to simulate obtaining information with different levels of precision. This scenario was composed by four subsets (D_1, D_2, D_3, and D_4) with 145 records each. However, they were simulated considering the standard deviation of the residual random error among incubation points of 2.0, 2.5, 3.0, and 3.5, respectively.

For all scenarios, each subset was simulated ten times in order to allow a more robust evaluation of the ability of different approaches to adjust in front of different scenarios.

The objective of the four scenarios were: (A) to evaluate circumstances where minimal sampling procedures are demanded; (B) to evaluated circumstances where an intense loss of information occurred; (C) to evaluate the robustness of the statistical approach regarding the loss of a specific sequential number of samples; and (D) to evaluate the robustness of the statistical approach when a low-precision dataset are available to estimate the parameters (Figure 1).
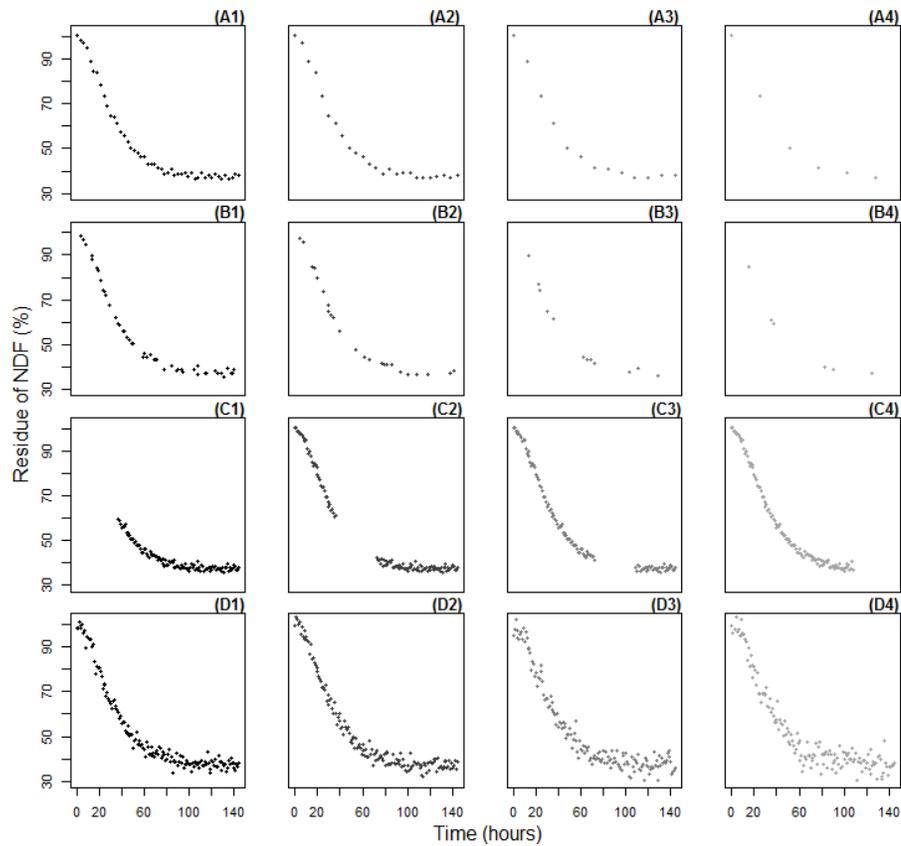
Figure 1 - Examples of the simulated scenarios: A, regular restriction in the number of incubations times; B, random loss of incubation times; C, loss of specific parts of the degradation curves; and D, variation in the precision of the incubation procedures.

2.3 Actual data

After the evaluation of simulated data, two real datasets were used to evaluate the Bayesian and frequentist approaches. Both datasets encompassed the evaluation of NDF degradation of a tropical grass (*Brachiaria decumbens*) in two experiments carried out according to complete 5×5 Latin square design with five treatments, which consisted of different supplementation schemes for growing cattle under grazing (FIGUEIRAS, 2013). The NDF degradation was measured within each experimental period (five for each experiment) according to the incubation times: 0, 3, 6, 9, 12, 24, 36, 48, 60, 72, 96, 120, and 144 hours. Therefore, ten different adjustments were performed (one for each treatment). In spite of few losses of bags (incubation points), each adjustment was performed, on average, with $n = 65$. Details of the treatments and incubation procedures can be found in Figueiras (2013).

2.4 Model fitting

Model fitting following the frequentist approach was performed according the ordinary least squares method with the solutions obtained by the Gauss-Newton algorithm using lme4 package from the R software (BATES *et al.*, 2015). Significance of the parameters was checked using the asymptotic confidence intervals.

For the fit according to Bayesian approach, the variance components were defined by $\tau = \dfrac{1}{\sigma^2} \sim Gam(10^{-3}, 10^{-3})$. Gamma and uniform distributions are standard choices of diffuse informative prior distributions to the variance components of model (GELMAN *et al.*, 2004). Considering that rumen kinetics data errors may not follow a normal distribution, the choice of priors distributions for the variance components was based on the construction of minimally informative priors. For model parameters was specifies prior distributions: $U, I, c \sim Gam\left(10^{-3}, 10^{-3}\right)$ where Gam denotes the gamma distribution (for which $Gam(a, b)$ has mean $= a/b$ and variance $= a/b^2$.

The *posterior* distribution was simulated by Monte Carlo Markov Chain (MCMC) process by using the statistical software OpenBugs (THOMAS *et al.*, 2016), which was interfaced with R via BRugs package (LIGGES, 2013) (LUNN *et al.*, 2009). Briefly, the *posterior* distribution is simulated by the MCMC method by constructing Markov Chains, for which partial distributions approximate the posterior density, and using Monte Carlo integration to compute integrals and expectations. The Monte Carlo error approximation is reduced by increasing the number of samples up to a point at which no further gain in accuracy is achieved in the approximation of the posterior summaries. Two chains with overdispersal initial values were determined for each parameter and chain mixing, auto-correlation, posterior distribution, and the Heildelberger-Welch diagnostic (HEIDELBERGER; WELCH, 1992) implemented via the BOA (Bayesian Output Analysis) package (SMITH, 2007). A minimum of 500,000 iterations was determined by simulation, of which 200,000 iterations composed a burn-in period, and after the burn-in period 300,000 iterations were saved to obtain posterior distribution estimates. The chains were thinned by a factor of 100, totalizing 3,000 simulated values in a final sample.

In the Bayesian approach, the *posterior* probabilities of the parameters were calculated as the proportion of chain iterations of the MCMC sampling method spent in each subset.

2.5 Evaluation of adjustment and comparison of parameter estimates

The quality of the adjustments was evaluated through the asymptotic standard deviation of the residual error (ASDR) using the likelihood estimator as follow:

$$ASDR = \sqrt{\dfrac{\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2}{n}}, \qquad (2)$$

where $\hat{Y}_i$ is the predict value, Y_i is the simulated or real (observed), and *n* is the number of records.

Specifically for the simulated dataset, the accuracy of the estimates for each parameter was evaluated by computing the bias (B). Such an assumption was based on the fact that estimates of the parameters obtaining by each method should converge to the parametric values used to simulate the scenarios and subsets. The average estimates of B (%) were calculated as:

$$B = \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}_i - \theta)}{\theta} \times 100, \qquad (3)$$

where $\hat{\theta}_i$ is the estimate of the parameter in the evaluation i for an specific scenario and subset ($i = 1, 2, \ldots, 10$), and θ is the parametric value.

The pattern of the average difference between the estimates of the parameters obtained by Bayesian and frequentist approaches was evaluated by using a paired t-test ($\alpha = 0.05$).

When necessary, descriptive plots were also drawn to a better understanding of the pattern of results.

The statistical computations was performed using the R statistical software (R CORE TEAM, 2017), except the MCMC method that was performed in the R statistical software (R CORE TEAM, 2017) assisted by statistical software WinBUGS (LUNN *et al.*, 2000).

3 Results

3.1 Simulated data

In the Bayesian simulation, the samples came from 300,000 iterations, already discounting the burn-in period, which resulted in 3,000 total record iterations. This total record number in each subset was verified for convergence of chains by the Heildelberger-Welch diagnostic. From this diagnostic, it was not registered any non-convergence in the subsets. Thus, 3,000 record iterations were sufficient to achieve convergence. Similarly, the frequentist approach did not show any case of unsuccessful convergence. During the model fitting, it was not determined any border for the parameters estimate, in order to evaluate if the approaches would be able to estimate the parameters in accordance with the biological coherence borders.

At first glance, both inferences showed a similar performance. On average, the parameters estimates, precision (Table 1) and bias, and the residual error (Table 2) were very similar to each other when both methods were numerically contrasted. It is important to note that for most subsets the sums of the potentially degradable and undegradable fractions ($U+I$, Table 1) were very close to 100%, which agreed with the biological and chemical assumptions of NDF degradation and added reliability to the model adjustments (Table 1). However, subset B_4, which was developed to test the highest level of random loss of incubation points, showed discrepant values for $U+I$ for both approaches (103.57 for Bayesian and 103.09 for frequentist). In a particular way, the adjustment of the subset B_4 through Bayesian approach lead to estimates of parameter c that did not match any expected biological pattern (Table 1) and with a highly prominent bias (Table 2).

Table 1 – Average estimates of the parameters potentially degradable fraction (U, %), common rate of lag and degradation (c, h^{-1}), undegradable fraction (I, %) obtained on simulated data using the Bayesian or frequentist approaches

Scenario[1]	Bayesian[2,3]				Frequentist[2,3]			
	U	c	I	$U+I$	U	c	I	$U+I$
A$_1$	63.00 (0.79)	0.058 (0.002)	36.99 (1.10)	99.99	62.97 (0.48)	0.058 (0.001)	37.06 (0.234)	100.04
A$_2$	63.11 (2.41)	0.065 (0.105)	36.15 (2.66)	99.27	62.78 (0.67)	0.059 (0.001)	37.19 (0.333)	99.97
A$_3$	62.96 (1.69)	0.071 (0.587)	37.21 (0.75)	100.17	63.05 (0.83)	0.059 (0.001)	37.22 (0.437)	100.27
A$_4$	62.69 (2.31)	0.061 (0.244)	37.09 (1.78)	99.77	62.70 (0.94)	0.059 (0.001)	37.24 (0.526)	99.94
B$_1$	63.30 (0.56)	0.059 (0.001)	37.08 (0.25)	100.38	63.30 (0.55)	0.059 (0.001)	37.08 (0.242)	100.39
B$_2$	63.53 (2.79)	0.169 (3.205)	35.16 (3.42)	98.69	63.40 (0.69)	0.059 (0.001)	37.05 (0.344)	100.46
B$_3$	62.44 (3.04)	0.063 (0.453)	36.98 (1.13)	99.42	62.58 (1.33)	0.057 (0.002)	36.99 (0.546)	99.57
B$_4$	67.05 (21.34)	1.131 (11.427)	36.52 (4.62)	103.57	65.74 (4.19)	0.060 (0.003)	37.35 (0.710)	103.09
C$_1$	62.93 (3.79)	0.058 (0.023)	36.93 (1.11)	99.85	62.94 (3.44)	0.058 (0.001)	37.02 (0.176)	99.96
C$_2$	63.12 (0.82)	0.058 (0.002)	36.90 (1.11)	100.02	63.03 (0.30)	0.059 (0.001)	37.06 (0.135)	100.09
C$_3$	63.05 (0.49)	0.058 (0.001)	37.01 (0.28)	100.06	63.05 (0.30)	0.058 (0.001)	37.01 (0.163)	100.07
C$_4$	62.98 (0.48)	0.059 (0.005)	37.10 (0.43)	100.08	62.97 (0.30)	0.059 (0.001)	37.11 (0.206)	100.09
D$_1$	63.24 (0.62)	0.059 (0.002)	36.95 (0.42)	100.18	63.24 (0.59)	0.059 (0.001)	36.96 (0.281)	100.19
D$_2$	63.18 (1.02)	0.058 (0.002)	36.76 (1.07)	99.94	63.11 (0.72)	0.059 (0.001)	36.93 (0.343)	100.04
D$_3$	63.26 (0.88)	0.059 (0.001)	36.98 (0.42)	100.25	63.27 (0.87)	0.059 (0.001)	36.99 (0.417)	100.26
D$_4$	63.04 (1.29)	0.059 (0.002)	37.08 (1.39)	100.12	62.97 (0.98)	0.060 (0.001)	37.28 (0.461)	100.25

[1] See text for details about scenarios.

[2] The values correspond to the average of ten adjustments for each subset within each scenario.

[3] The values between parentheses correspond to the standard deviation among adjusted models ($n = 10$).

In spite of being very similar, the estimates of the undegradable fraction I differed between Bayesian and frequentist approaches ($P<0.05$). On average, I fraction was higher for the frequentist approach (Table 3). Is must be emphasized that this pattern was kept even when not considering subset B_4 in the paired t-test (data not shown). There was no difference between approaches with regard parameters U and c ($P \geq 0.16$).

Table 2 – Asymptotic standard deviation of residual error (ASDR) for the adjusted model and average bias (B, %) for the estimates of the parameters potentially degradable fraction (U, %), common rate of lag and degradation (c, h^{-1}), undergradable fraction (I, %) obtained on simulated data using the Bayesian or frequentist approaches

Scenario[1]	Bayesian[2,3]				Frequentist[2,3]			
	ASDR	B			ASDR	B		
		U	c	I		U	c	I
A_1	0.982	0.130 (0.47)	-0.380 (0.784)	-0.237 (0.439)	0.982	0.085 (0.480)	-0.246 (0.867)	-0.043 (0.547)
A_2	2.617	0.303 (1.54)	5.716 (15.932)	-3.331 (10.553)	0.850	-0.218 (0.626)	0.081 (0.885)	0.286 (0.443)
A_3	2.402	0.067 (1.77)	7.093 (21.490)	0.339 (0.703)	1.007	0.202 (1.707)	0.237 (1.623)	0.367 (0.707)
A_4	1.363	-0.370 (2.35)	2.712 (5.450)	-0.004 (1.689)	1.158	-0.341 (2.169)	0.269 (1.180)	0.420 (0.938)
B_1	1.150	0.602 (0.73)	0.478 (1.270)	-0.0002 (0.573)	1.150	0.610 (0.734)	0.485 (1.272)	0.006 (0.570)
B_2	6.547	0.963 (2.14)	18.539 (32.591)	-10.255 (31.930)	1.321	0.767 (0.670)	0.567 (1.648)	-0.076 (0.800)
B_3	6.130	-0.763 (2.24)	5.469 (13.104)	-0.285 (1.375)	4.456	-0.536 (2.228)	-2.055 (2.860)	-0.265 (1.341)
B_4	4.520	6.562 (10.62)	29.10 (35.139)	-1.722 (4.842)	0.974	4.485 (7.483)	2.912 (4.628)	0.665 (2.525)
C_1	1.177	0.014 (6.61)	-0.664 (3.996)	-0.422 (0.730)	1.163	0.028 (6.608)	-0.544 (3.872)	-0.161 (0.603)
C_2	1.243	0.324 (0.53)	-0.146 (1.247)	-0.503 (1.307)	1.205	0.181 (0.307)	0.173 (0.829)	-0.059 (0.347)
C_3	1.179	0.200 (0.25)	-0.050 (0.513)	-0.190 (0.328)	1.180	0.209 (0.255)	-0.039 (0.510)	-0.175 (0.323)
C_4	1.173	0.092 (0.34)	0.226 (0.675)	0.063 (0.450)	1.173	0.088 (0.332)	0.199 (0.546)	0.090 (0.393)
D_1	1.852	0.507 (0.89)	0.206 (1.684)	-0.370 (0.923)	1.851	0.506 (0.886)	0.219 (1.708)	-0.339 (0.929)
D_2	2.583	0.410 (1.07)	-0.245 (2.756)	-0.892 (1.842)	2.569	0.305 (1.030)	0.083 (2.367)	-0.408 (0.869)
D_3	2.806	0.545 (1.38)	0.066 (1.595)	-0.273 (1.114)	2.805	0.556 (1.378)	0.066 (1.594)	-0.253 (1.108)
D_4	3.624	0.194 (0.84)	1.209 (2.892)	-0.010 (1.312)	3.642	0.083 (0.873)	1.572 (2.700)	0.513 (1.150)

[1] See text for details about scenarios.
[2] The values correspond to the averages of ten adjustments for each subset within each scenario.
[3] The values between parentheses correspond to standard deviation of bias among ten adjustments.

Table 3 - Paired t-test to compare the estimates of the parameters potentially degradable fraction (U, %), common rate of lag and degradation (c, h^{-1}), undegradable fraction (I, %) obtained using Bayesian or frequentist approaches for simulated and actual dataset

Parameters	Bayesian	Frequentist		t-test	
	Mean		n	s($\overline{d}$)	P value
Simulated values					
U	63.33±2.89	63.20±1.09		0.396	0.226
c	0.069±0.346	0.059±0.001	16	0.027	0.166
I	36.80±1.38	37.10±0.34		0.533	0.040
Real values					
U	68.51±4.30	66.917±3.57		3.859	0.223
c	0.051±0.045	0.051±0.009	10	0.003	0.919
I	27.50±4.18	29.46±3.28		4.452	0.197

3.2 Actual dataset

In general, the pattern of the results regarding $U+I$ obtained for the actual dataset followed what has been observed for the simulated dataset (Table 4). There was little variation among treatments within each experiment for the frequentist approach, and for Experiment 2 when Bayesian approach was applied. However, the Bayesian approach tended to over and underestimate the fractions U and I for treatments 2 and 5 within Experiment 1, respectively (Table 4), when compared with the other treatments and with the estimates obtained with frequentist approach. It must be pointed out that those treatments were the ones with the lowest precision among all evaluated treatments.

In spite of that behavior, there were no differences between approaches with regard the parameter estimates (P≥0.19, Table 3). Nonetheless, similarly to the simulated dataset, the average estimates of parameter I were numerically higher for the frequentist approach (Table 3). The relevance of this pattern will be discussed with more details in the next section.

Table 4 - Estimates of the parameters potentially degradable fraction (U, %), common rate of lag and degradation (c, h^{-1}), undegradable fraction (I, %) obtained on actual data (two experiments and five treatments per experiment) using Bayesian or frequentist approaches

Real data	Bayesian					Frequentist				
	U	c	I	$U+I$	ASDR[1]	U	c	I	$U+I$	ASDR
Exp. 1										
T_1	69.09±3.94	0.042±0.005	28.02±3.95	97.12	9.691	68.60±3.55	0.043±0.004	28.75±3.34	97.35	9.686
T_2	76.96±5.07	0.043±0.379	19.45±5.23	96.41	10.309	75.66±3.97	0.037±0.003	21.09±3.94	96.75	8.576
T_3	71.77±4.62	0.035±0.004	27.99±4.88	99.76	9.248	70.85±3.88	0.036±0.003	29.16±3.88	100.01	9.239
T_4	77.76±4.15	0.038±0.004	25.53±4.20	103.29	9.408	77.10±3.74	0.038±0.003	26.42±3.67	103.52	9.403
T_5	83.22±11.73	0.027±0.029	11.58±12.52	94.80	15.555	70.72±7.87	0.033±0.006	26.12±8.06	96.83	15.120
Exp. 2										
T_1	58.19±2.84	0.060±0.006	35.37±2.53	93.57	8.394	58.14±2.69	0.060±0.005	35.54±2.17	93.69	8.392
T_2	61.50±2.60	0.063±0.005	32.34±2.04	93.85	7.297	61.51±2.32	0.063±0.005	32.46±1.83	93.97	7.295
T_3	64.72±2.84	0.072±0.007	30.65±2.12	95.38	8.346	64.73±2.61	0.072±0.058	30.76±1.95	95.49	8.345
T_4	60.52±2.87	0.070±0.007	33.52±2.32	94.05	8.890	60.51±2.79	0.070±0.006	33.65±2.10	94.16	8.889
T_5	61.38±2.36	0.060±0.005	30.51±2.00	91.90	7.138	61.35±2.29	0.060±0.004	30.63±1.85	91.98	7.137

[1] Asymptotic standard deviation of residual error (ASDR) for the adjusted model.

4 Discussion

The essence of the Bayesian method is there is no logical distinction between model parameters and data (BEAUMONT and RANNALA, 2004). Both of them are random variables with a joint probability distribution that is specified by each evaluated model. The aim of the Bayesian inference is to calculate the posterior distribution of the parameters, the conditional distribution of parameters given a data, assisted by a prior distribution. Conversely, the frequentist approach is based on maximizing the probability of the data given the parameters (that is, maximizing the likelihood as a function of the parameter for a fixed dataset).

In fact, both inferences aim to estimate the parameters with the least possible error. The Bayesian inference has a specific advantage comparing to frequentist inference that is to consider the *prior* information (SILVA *et al.*, 2011; 2013). However, it is common in practice to utilize a default or objective prior distribution. Then, the Bayes theorem would not provide any guarantee as to performance (BAYARRI and BERGER, 2004). Another problem is the choice of an improper prior distribution. In this case, the Bayes theorem may generate an incorrect performance and, consequently, an improper posterior distribution.

In terms of mathematic, the results coming from frequentist inference and Bayesian inference with non-informative priors would tend to be the same. Browne and Draper (2006) compared the Bayesian approach (without prior influence) and frequentist approach (likelihood-based) in different types of models and concluded that both methods lead to similar and unbiased estimates. Such a conclusion agrees, in general, with the results obtained for the scenarios A, C, and D, where both approaches seems to present similar robustness in front of an organized minimal sampling (Scenario A); an loss of an entire section of the curve, but preserving the other parts (Scenario C); and a dataset with increased random error (Scenario D).

However, the choice of non-informative prior distribution can affect the inferences; noticeably in cases where the number of records/groups evaluated are small or when the variance of the records/group is close to zero (GELMAN, 2004; 2006). For instance, the uniform distribution is commonly considered an improper prior distribution for the variance parameter. The uniform distribution $(0, A)$ (*i.e.* uniform distribution with values in the range 0 and A) can produce a limited proper posterior distribution as $A \to \infty$, as long as the number of records/groups. Thus, for a finite but sufficiently large A, the inferences are not sensitive to the choice of A (GELMAN, 2006). In order to evaluate both inferences as balanced as possible in this paper, it was used a "weakly informative prior" gamma distribution (U, I, c ~Gam $(10^{-3}, 10^{-3})$) for the parameters. The choice of gamma distribution was based on the fact that only positive values belongs to this distribution and none parameter of the evaluated model can assume negative values (Equation 1).

However, the constraint arisen from limit data availability along with using a non-informative prior distribution seemed to affect the estimates for subset B_4, leading to great bias of the estimates of incubation residue (Figure 2) and estimates of degradation rate without a biological coherence (Tables 1 and 2). It should be noted that a similar pattern was not observed for frequentist approach (Figure 3). When random losses are simulated, it is not possible to assure where the remained incubation points would be located regarding time. Therefore, it is possible to happen that points may be located very close to each other,

which could cause a lack of agreement between the real information and the *prior* distribution used here. Therefore, it may suggest that under a high level of random loss of incubations points, a more detailed study of the *prior* distribution (maybe coming from information from other studies or meta-analyses) must be done to assure reliability and accuracy of predicted values. It must be highlighted that utilization of Bayesian approach to rumen kinetics is not a routine in nutritional studies and knowledge on distribution of the related parameters is not available yet.
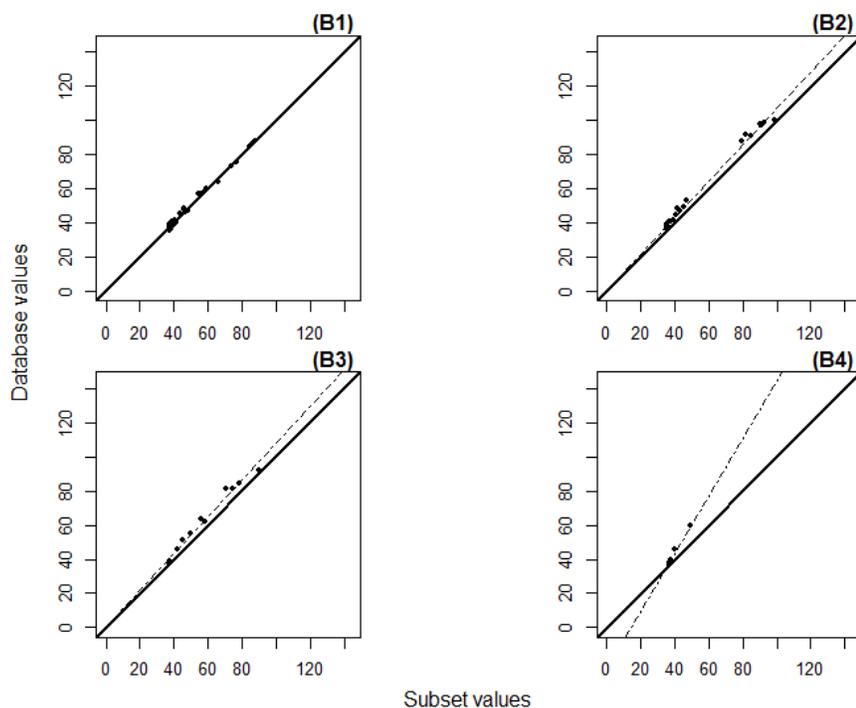

Figure 2 - Descriptive relationship between actual (from subsets) and predicted according to Bayesian approach (from model adjustment) of the NDF residues as a function of time for scenario B. The solid lines represent the equality line.
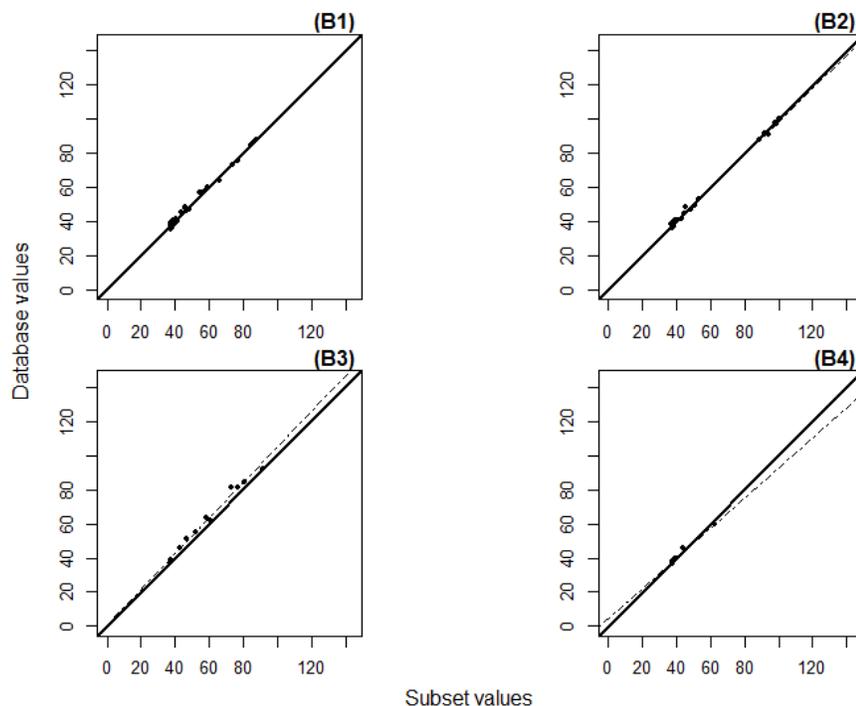
Figure 3 - Descriptive relationship between actual (from subsets) and predicted according to frequentist approach (from model fitted) of the NDF residues as a function of time for scenario B. The solid lines represent the equality line.

In spite of the previous discussion, one constraint was observed for Bayesian approach. The evaluation of simulated dataset showed a systematic underestimation of undegradable fraction of NDF (Tables 1, 2, and 3). The dimension of fractions *U* and *I* are inherent to the feed and cannot be influenced by dietary characteristics (DETMANN *et al.*, 2008). In other words, when a diet favors or disfavors the ruminal degradation of fiber, its effects would be perceived only on degradation rate. In spite of a non-significant difference between the approaches evaluated here, the evaluation of the real dataset showed numerically lower estimates of fraction *I* when the Bayesian approach was used (Tables 3 and 4). The fraction *I* was evaluated in both experiments by Figueiras (2013) using a long-term incubation procedure (288 hours) as recommended by Valente *et al.* (2011). This procedure is used based on the minimal time necessary to incubation residue becomes statistically similar to the residue theoretically obtained at infinite time. Figueiras (2013) found estimates for *I* fraction of 29.68 and 33.76% in Experiments 1 and 2, respectively. From this, it is noted that frequentist approach produced estimated closer to that values (26.31 and 32.61%) when compared to Bayesian approach (22.51 and 32.48%). Moreover, the high variation among treatments for the estimates of *U* and *I* fractions in Experiment 1 (Table 4) indicates problems with regard Bayesian approach. The forage evaluated within each experiment was the same. Therefore, variations between treatments must be only minimal and random as the dimensions of the fractions are inherent to the feed itself and cannot vary according to different supplements.

Mertens (2005) highlighted that the final points of an incubation procedure should provide information that allow a good adjustment of the model regarding asymptote. In other words, at least the final incubation point should be located on a specific time in order to give reliable information of the dimension of undegradable fraction and allow a more accurate adjustment. The endpoint at 144 hours seemed a good procedure when the frequentist approach was used, but not when Bayesian approach was applied. It may be a reflex of the lack of a more reliable *prior* information, as discussed before.

Conclusions

The Bayesian approach and frequentist approach presented reliable estimates for the majority of the data tested. Therefore, in specific cases with short random records number, the Bayesian approach showed greater bias of the incubation residue estimates and estimates of degradation rate without a biological coherence, compared to frequentist inference. In another words, the Bayesian approach fitted with non-informative prior, presented less flexible than frequentist inference. Nevertheless, it is emphasized the importance of the background information before the modeling, especially in a Bayesian approach, in order to define proper prior distributions. Future thorough studies about the influence of non-informative prior correlated to parameters estimated are necessary.

Acknowledgments

- BONFÁ, H. C.; DETMANN, E.; SILVA, F. F.; FIGUEIRAS, J. F. Comparação da abordagem Bayesiana e Frequentista para ajuste de modelos não-lineares: uma aplicação ao modelo gama-tempo-dependente para descrição degradação da fibra em detergente neutro. *Rev. Bras. Biom.*, Lavras, v.36, n.3, p.625-648, 2018.

- *RESUMO: Objetivou-se avaliar e comparar a eficiência das inferências Bayesiana e Frequentista frente a descrição da degradação ruminal da fibra em detergente neutro utilizando um modelo do tipo gama-tempo-dependente. Um banco de dados simulado composto por quatro cenários foi utilizado: restrição regular do número de tempos de incubação, perda aleatória de tempos de incubação, perda de partes específicas das curvas de degradação, variação na precisão dos procedimentos de incubação. Os cenários foram divididos em quatro subconjuntos cada, os subconjuntos foram simulados dez vezes, permitindo uma ampla avaliação da capacidade de ajuste das diferentes abordagens frente os diferentes cenários. Dois bancos de dados reais foram utilizados para avaliar as abordagens, os dados reais foram compostos da degradação da NDF de uma gramínea tropical (Brachiaria decumbes). O modelo foi ajustado de acordo com as características de cada abordagem e comparados via gráficos e avaliadores. A abordagem bayesiana e a abordagem Frequentista apresentaram estimativas confiáveis para a maioria dos dados avaliados. Entretanto, a abordagem bayesiana apresentou um grande viés das estimativas do resíduo e das estimativas da taxa de degradação de incubação, não apresentando uma coerência biológica para os parâmetros, em comparação com a abordagem Frequentista. Ou seja, a abordagem Bayesiana ajustada com priores não informativas, apresentou-se menos flexível do*

que a abordagem Frequentista. No entanto, ressalta-se a importância das informações a priori, especialmente para a abordagem bayesiana, a fim de definir priores apropriadas ao modelo. Estudos futuros mais aprofundados sobra a influência das priores não informativas sobre a estimativa de parâmetros são necessários.

- *PALAVRAS-CHAVE: Gama-tempo-dependente; inferência; degradação ruminal; método in situ.*

References

BATES, D.; MACHLER, M.; BOLKER, B. M.; WALKER, S. C. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software,* v.67, n.1, p.1-48, 2015.

BAYARRI, M. J.; BERGER, J. O. The interplay of Bayesian and frequentist analysis. *Statistical Science,* v.19, n.1, p.58-80, 2004.

BEAUMONT, M. A.; RANNALA, B. The Bayesian revolution in genetics, *Nature Reviews Genetics,* v.5, p.251-261, 2004.

BROWNE, W. J.; DRAPER, D. A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis,* v.1, n.3, p.473-514, 2006.

DETMANN, E.; CECON, P. R.; PAULINO, M. F.; ZERVOUDAKIS, J. T.; VALADARES FILHO, S. C.; ARAÚJO, C. V. Estimação de parâmetros de cinética de trânsito de partículas em bovinos sob pastejo por diferentes sequências amostrais. *Revista Brasileira de Zootecnia,* v.30, p.222-230, 2001.

DETMANN, E.; PAULINO, M.F.; VALADARES FILHO, S.C. Avaliação nutricional de alimentos ou de dietas? Uma abordagem conceitual. In: SIMPÓSIO DE PRODUÇÃO DE GADO DE CORTE, 6, 2008, Viçosa. *Proceedings…* Viçosa: DZO-UFV, 2008. p.21-52.

ELLIS, W. C.; MATIS, J. H.; HILL, T. M.; MURPHY, M. R. Methodology for estimating digestion and passage kinetics of forages. In: FAHEY JR., G. C. (Eds.) *Forage quality, evaluation, and utilization.* Madison: American Society of Agronomy, 1994. p.682-756.

FIGUEIRAS, J. F. *Desempenho nutricional de bovinos em pastejo suplementados durante os períodos de transição seca-águas e de águas.* 2013. 73p. Thesis (Ph.D.) – Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, 2013.

GELFAND, A. E.; HILLS, S.; RACINE POON, A.; SMITH, A. E. M. Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association,* v.85, p.972-985, 1990.

GELMAN, A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis,* v.1, n.3, p.515-533, 2006.

GELMAN, A.; CARLIN, J.; STERN, H.; RUBIN, D. *Bayesian data analysis.* 3.ed. Texts in Statistical Science. London: Chapman and Hall, 2004. 656p.

HEIDELBERGER, P.; WELCH, P. Simulation run length control in the presence of an initial transient, *Operations Research*, v.31, p.1109–114, 1992.

HUHTANEN, P.; AHVENJÄRVI, S.; BRODERICK, G. A.; REYNAL, S. M.; SHINGFIELD, K. J. Quantifying ruminal digestion of organic matter and neutral detergent fiber using the omasal sampling technique in cattle - a meta-analysis. *Journal of Dairy Science*, v.93, p.3203-3215, 2010.

JEFFREYS, H. *Theory of probability.* Oxford: Claredon, 1961. 325p.

LESAFFRE, E.; LAWSON, A. B. *Bayesian Biostatistics.* Chichester: John Wiley & Sons, 2012. 516p.

LIGGES, U. *BRugs: R interface to the OpenBUGS MCMC software*. R package version 0.8-3. Acessed May. 24, 2017. http;//cran.r-project.org/web/packages/BRugs.

LUNN, D. J.; SPIEGELHALTER, D.; THOMAS, A.; BEST, N. The BUGS project: evolution, critique, and future directions. *Statistics in Medicine,* v.28, p.3049-3067. 2009.

MERTENS, D. R. Rate and extent of digestion. In: DIJKSTRA, J.; FORBES, J. M.; FRANCE, J. (Eds.) *Quantitative aspects of ruminant digestion and metabolism.* 2.ed. Wallingford: CAB International, 2005. p.13-47.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org, 2017.

ROSSI, M. R.; MARTINS, E. N.; GUEDES, T. A.; JOBIM, C. C. Bayesian analysis for comparison of nonlinear regression model parameters: an application to ruminal degradability data. *Revista Brasileira de Zootecnia,* v.39, n.2, p.419-424, 2010.

SILVA, F. F.; SAFADI, T.; MUNIZ, J. A.; ROSA, G. J. M.; AQUINO, L. H.; MOURÃO, G. B.; SILVA, C. H. O. Bayesian analysis of autoregressive panel data model: application in genetic evaluation of beef cattle. *Scientia Agricola,* v.68, n.2, p.237-245, 2011.

SILVA, F. F.; VIANA, J. M. S.; FARIA, V. R.; RESENDE, M. D. V. Bayesian inference of mixed models in quantitative genetics of crop species. *Theoretical and Applied Genetics,* v.126, p.1749-1761, 2013.

SMITH, B.J. boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference. *Journal of Statistical Software.,* v.21, p.1-37, 2007.

THOMAS, A.; O'HARA, B.; LIGGES, U.; STURTZ, S. Making BUGS Open. *R News,* v.6, n.1, p.12-17, 2006.

VALENTE, T. N. P.; DETMANN, E.; QUEIROZ, A. C.; VALADARES FILHO, S. C.; GOMES, D. I.; FIGUEIRAS, J. F. Evaluation of ruminal degradation profiles of forages using bags made from different textiles. *Revista Brasileira de Zootecnia,* v.40, n.11, p.2565-2573, 2011.

VAN MILGEN, J.; MURPHY, L. L.; BERGER, L. L. A compartmental model to analyse ruminal digestion. *Journal of Dairy Science,* v.74, p.2515-2529, 1991.

Appendix

Code for R Software.
BAYESIANA AND FREQUENTISTA APPROACH:
APPLYING TO A GAMMA-TIME-DEPENDENT MODEL

```
# Determinining library
Dir <- setwd("C:/Users/ ")

# Simulation #
  # Van Milgen et al., (1991) -> y = beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error
  # Values: Figueiras (2013) Thesis #

for(i in 1:10)                              # Loop 1
  {
beta = 62.92
sigma = 37.08
lambda = 0.059
time = seq(0, 144, 1)
error = rnorm(mean=0, sd=1, n=length(time))
y = beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error

Y=as.data.frame(rbind(y,time))
Yt = as.data.frame(t(Y))

# Graphic #
plot(Yt$y ~ Yt$time)

# Building scenarios #
# A = Regular restriction in the number of incubation times #
  # A1 = 3-3 hours -> n=48
  # A2 = 6-6 hours -> n=24
  # A3 = 12-12 hours -> n=12
  # A4 = 24-24 hours -> n=6

A1 = Yt[seq(1,145,3),]
A2 = Yt[seq(1,145,6),]
A3 = Yt[seq(1,145,12),]
A4 = Yt[seq(1,145,24),]

par(mfrow=c(2,2),mar=c(4,4,1.6,0.5), pty="s")
plot(A1$y ~ A1$time, xlim=range(0,144),ylim=range(35,100), col="blue", pch=20)
plot(A2$y ~ A2$time, xlim=range(0,144),ylim=range(35,100), col="red", pch=20)
plot(A3$y ~ A3$time, xlim=range(0,144),ylim=range(35,100), col="green", pch=20)
plot(A4$y ~ A4$time, xlim=range(0,144),ylim=range(35,100), col="purple", pch=20)

# B = Random loss of incubation times #
```

```
# B1 = 33% of total n -> n=48
# B2 = 16% of total n -> n=24
# B3 = 8% of total n -> n=12
# B4 = 4% of total n -> n=6

B1 = Yt[sample(1:145, 48,replace=F),]
B2 = Yt[sample(1:145, 24,replace=F),]
B3 = Yt[sample(1:145, 12,replace=F),]
B4 = Yt[sample(1:145, 6,replace=F),]

par(mfrow=c(2,2),mar=c(4,4,1.6,0.5), pty="s")
plot(B1$y ~ B1$time, xlim=range(0,144),ylim=range(35,100), col="blue", pch=20)
plot(B2$y ~ B2$time, xlim=range(0,144),ylim=range(35,100), col="red", pch=20)
plot(B3$y ~ B3$time, xlim=range(0,144),ylim=range(35,100), col="green", pch=20)
plot(B4$y ~ B4$time, xlim=range(0,144),ylim=range(35,100), col="purple", pch=20)

# C = Loss of specific parts of the degradation curves #
# C1 = 75% of total n; 2th,3th and 4th part -> n=108
# C2 = 75% of total n; 1th,3th and 4th part -> n=108
# C3 = 75% of total n; 1th,2th and 4th part -> n=108
# C4 = 75% of total n; 1th,2th and 3th part -> n=108

C1 = subset(Yt, time>=37)
C2 = subset(Yt, 36>=time | time>=73)
C3 = subset(Yt, 72>=time | time>=109)
C4 = subset(Yt, 108>=time)

par(mfrow=c(2,2),mar=c(4,4,1.6,0.5), pty="s")
plot(C1$y ~ C1$time, xlim=range(0,144),ylim=range(35,100), col="blue", pch=20)
plot(C2$y ~ C2$time, xlim=range(0,144),ylim=range(35,100), col="red", pch=20)
plot(C3$y ~ C3$time, xlim=range(0,144),ylim=range(35,100), col="green", pch=20)
plot(C4$y ~ C4$time, xlim=range(0,144),ylim=range(35,100), col="purple", pch=20)

# D = Variation in the precision of the incubation procedures #
# D1 = error: sd=2    -> n=144
# D2 = error: sd=2.5  -> n=144
# D3 = error: sd=3    -> n=144
# D4 = error: sd=3.5  -> n=144

error1 = rnorm(mean=0, sd=2, n=length(time))
error2 = rnorm(mean=0, sd=2.5, n=length(time))
error3 = rnorm(mean=0, sd=3, n=length(time))
error4 = rnorm(mean=0, sd=3.5, n=length(time))

D1 = as.data.frame(t(rbind(y=beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error1,time)))
D2 = as.data.frame(t(rbind(y=beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error2,time)))
D3 = as.data.frame(t(rbind(y=beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error3,time)))
D4 = as.data.frame(t(rbind(y=beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error4,time)))

par(mfrow=c(2,2),mar=c(4,4,1.6,0.5), pty="s")
plot(D1$y ~ D1$time, xlim=range(0,144),ylim=range(35,100), col="blue", pch=20)
```

```
plot(D2$y ~ D2$time, xlim=range(0,144),ylim=range(35,100), col="red", pch=20)
plot(D3$y ~ D3$time, xlim=range(0,144),ylim=range(35,100), col="green", pch=20)
plot(D4$y ~ D4$time, xlim=range(0,144),ylim=range(35,100), col="purple", pch=20)


# Merging Datas #
library(plyr)
Yc=join_all(list(Yt,A1,A2,A3,A4,B1,B2,B3,B4,C1,C2,C3,C4,D1,D2,D3,D4),by="time")
colnames(Yc)[1:18]=c("Y","time","A1","A2","A3","A4","B1","B2","B3","B4","C1","C2","C3","C4","D1","
D2","D3","D4")


# Plot with all scenarios #

par(mfrow=c(4,4), mar = c(0,0,1.4,1),oma = c(3.6,3,0,0),  pty="s")
# Scenario A #
plot(A1$y ~  A1$time, xlim=range(0,144), xaxt='n', ylim=range(30,102), col="gray0", pch=20, ylab="",
xlab="", main="                     (A1)")
plot(A2$y ~  A2$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102),    col="gray25", pch=20,
ylab="", xlab="", main="                     (A2)")
plot(A3$y ~  A3$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray50", pch=20,
ylab="", xlab="", main="                     (A3)")
plot(A4$y ~  A4$time, xlim=range(0,135),yaxt='n', xaxt='n',ylim=range(30,102), col="gray65", pch=20,
ylab="", xlab="", main="                        (A4)")
# Scenario B #
plot(B1$y ~ B1$time, xlim=range(0,144),xaxt='n',ylim=range(30,102), col="gray0", pch=20, ylab="", xlab="",
main="                 (B1)")
plot(B2$y ~  B2$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray25", pch=20,
ylab="", xlab="", main="                        (B2)")
plot(B3$y ~  B3$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray50", pch=20,
ylab="", xlab="", main="                        (B3)")
plot(B4$y ~  B4$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray65", pch=20,
ylab="", xlab="", main="                     (B4)")
# Scenario C #
plot(C1$y ~ C1$time, xlim=range(0,144),xaxt='n',ylim=range(30,102), col="gray0", pch=20, ylab="", xlab="",
main="                 (C1)")
plot(C2$y ~  C2$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray25", pch=20,
ylab="", xlab="", main="                        (C2)")
plot(C3$y ~  C3$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray50", pch=20,
ylab="", xlab="", main="                        (C3)")
plot(C4$y ~  C4$time, xlim=range(0,144),yaxt='n', xaxt='n',ylim=range(30,102), col="gray65", pch=20,
ylab="", xlab="", main="                     (C4)")
# Scenario D #
plot(D1$y ~ D1$time, xlim=range(0,144),ylim=range(30,102), col="gray0", pch=20, ylab="", xlab="", main="
(D1)")
mtext("                                                                Time (hours)",side=1,
line=2.5)
mtext("                                                   Residue of NDF (%)",side=2,
line=2.5)
plot(D2$y ~  D2$time, xlim=range(0,144),yaxt='n',ylim=range(30,102), col="gray25", pch=20, ylab="",
xlab="", main="                  (D2)")
plot(D3$y ~  D3$time, xlim=range(0,144),yaxt='n',ylim=range(30,102), col="gray50", pch=20, ylab="",
xlab="", main="                  (D3)")
```

```
plot(D4$y ~ D4$time, xlim=range(0,144),yaxt='n',ylim=range(30,102), col="gray65", pch=20, ylab="",
xlab="", main="                    (D4)")
```


```
###############################
# Modeling - Bayesian Model #
###############################

# Library load #
library(BRugs)

# Start of the loop
# Selecting Dataset/Subset
for(data in c("A1","A2","A3","A4","B1","B2","B3","B4","C1","C2","C3","C4","D1","D2","D3","D4" ))     #
Loop 2
  {
eval(parse(text=paste("dat=",data,sep="")))

# Van Milgen et al., (1991) Adapted -> y = beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error

# Building the dataset #
dd <- list(y=dat$y,
        x=dat$time,
        N=nrow(dat))

bugsData(dd , fileName = file.path(getwd(), "dataDeg.txt"))

# Creating the inits #
inits <- list(
  list(tau=runif(1, 0, 5), lambda=runif(1, 0, 5), sigma=runif(1, 0, 5),
         beta=runif(1, 0, 5)),

  list(tau=runif(1, 0, 5), lambda=runif(1, 0, 5), sigma=runif(1, 0, 5),
         beta=runif(1, 0, 5)))

bugsInits(inits, numChains=2, fileName=c(file.path(getwd(), "Inits1.txt"),
                        file.path(getwd(), "Inits2.txt")))

# Building the Model #
Degmodel <- function()
{
  for(i in 1:N)                              # Loop 3
   {
   y[i] ~ dnorm(mu[i], tau)
   mu[i] <- beta*(1+lambda*x[i])*exp(-lambda*x[i])+sigma
  }
 sigma  ~ dgamma(0.001, 0.001)   # non-informative gamma distribuction
 lambda ~ dgamma(0.001, 0.001)   # non-informative gamma distribuction
 beta   ~ dgamma(0.001, 0.001)   # non-informative gamma distribuction
 tau    ~ dgamma(0.001, 0.001)   # non-informative gamma distribuction
 }                                  # Loop 3 ends
```

```
writeModel(Degmodel, file.path(getwd(), "DegModel.txt"))

# Bayesian using Open Bugs #
modelCheck("DegModel.txt")              # check model file
modelData("dataDeg.txt")                # read data file
modelCompile(numChains=2)               # compile model with 2 chains
modelInits(c("Inits1.txt","Inits2.txt"))    # read init data file

modelUpdate(200000)                     # burn in
samplesSet(c("beta", "lambda", "sigma"))      # parameters monitored
modelUpdate(300000, thin=100)           # interations

# Sumarized results
# samplesStats("*")
eval(parse(text=paste("Stats_",data,"<- samplesStats('*')",sep="")))

# Export the iterations data's
samplesCoda("beta",paste(Dir,"/beta/",sep=""))
samplesCoda("lambda",paste(Dir,"/lambda/",sep=""))
samplesCoda("sigma",paste(Dir,"/sigma/",sep=""))

# Graphics #
# samplesHistory("beta")
# samplesHistory("lambda")
# samplesHistory("sigma")
# samplesHistory("*", mfrow = c(3,1))

# Library load #
library(boa)

# Building data's to Convergence Diagnostics
beta1 <- read.table("beta/CODAchain1.txt", sep="")
beta2 <- read.table("beta/CODAchain2.txt", sep="")
lambda1 <- read.table("lambda/CODAchain1.txt", sep="")
lambda2 <- read.table("lambda/CODAchain2.txt", sep="")
sigma1 <- read.table("sigma/CODAchain1.txt", sep="")
sigma2 <- read.table("sigma/CODAchain2.txt", sep="")

coda <- as.matrix(data.frame(beta1[,2],beta2[,2],lambda1[,2],lambda2[,2],sigma1[,2],sigma2[,2]))   # Create a
matrix data
colnames(coda)[1:6] = paste(c("beta1","beta2","lambda1","lambda2","sigma1","sigma2"),data,sep="_")   #
Rename the col's
rownames(coda)=beta1[,1]                                         # Rename the row's

# Convergence Diagnostics - Heildelberger and Welch (1992) (error=0.1, alpha=0.05) #
eval(parse(text=paste("Diag_",data,"<- boa.handw(coda,0.1,0.05)",sep="")))

# Asymptotic standard deviation of the residual error (ASDR) #
y_b      =      (samplesStats('*')[[1]][1])*(1      +      (samplesStats('*')[[1]][2])*dat$time)*exp(-
(samplesStats('*')[[1]][2])*dat$time) + (samplesStats('*')[[1]][3])
```

```
ASDR <- eval(parse(text=paste("ASDR_",data,"<- (sum((dat$y-y_b)^2)/(length(dat$y)))^0.5",sep="")))

# Bias (B) #
V_1<-eval(parse(text=paste("V_beta_", data,
    "<- (((samplesStats('*')[[1]][1])-beta)/(length(dat$y)))",sep="")))     # Bias beta
V_2<-eval(parse(text=paste("V_lambda_", data,
    "<- (((samplesStats('*')[[1]][2])-lambda)/(length(dat$y)))",sep="")))   # Bias lambda
V_3<-eval(parse(text=paste("V_sigma_", data,
    "<- (((samplesStats('*')[[1]][3])-sigma)/(length(dat$y)))",sep="")))    # Bias sigma

# Assessors (As) #
As <- data.frame(ASDR,V_1,V_2,V_3)
colnames(As)[1:4]=c("ASDR","Bias_beta", "Bias_lambda", "Bias_sigma")

eval(parse(text=paste("As_", data, " <- As",sep="")))

#################################
# Modeling - Frequentist Model #
#################################

# Van Milgen et al., (1991) Adapted -> y = beta*(1 + lambda*time)*exp(-lambda*time) + sigma + error

# Library load #
library(lme4)

# Renaming the parameters #
beta_=beta
lambda_=lambda
sigma_=sigma
t=dat$time

# Building the model #
start = c(beta_=beta, lambda_=lambda, sigma_=sigma)
model = ~ beta_*(1 + lambda_*t)*exp(-lambda_*t) + sigma_
model_d = deriv(model, namevec=c("beta_", "lambda_", "sigma_"),
        function.arg=c("input","beta_", "lambda_", "sigma_"))

M_y = nls(y ~ model_d(time, beta_, lambda_, sigma_), data=dat, start=start)

eval(parse(text=paste("Stats2_", data, "<- coef(summary(M_y))",sep="")))

# Asymptotic standard deviation of the residual error (ASDR) #
ASDR2               <-               eval(parse(text=paste("ASDR_",data,"<-               (sum((dat-
predict(M_y))^2)/(length(dat$y)))^0.5",sep="")))

# Bias (B) #
V2_1 <- eval(parse(text=paste("V2_beta_", data,
                "<- (((coef(summary(M_y))[[1]])-beta)/(length(dat$y)))",sep="")))     # Bias beta
V2_2 <- eval(parse(text=paste("V2_lambda_", data,
                "<- (((coef(summary(M_y))[[2]])-lambda)/(length(dat$y)))",sep="")))   # Bias lambda
```

```
V2_3 <- eval(parse(text=paste("V2_sigma_", data,
                 "<- (((coef(summary(M_y))[[3]])-sigma)/(length(dat$y)))",sep="")))    # Bias sigma

# Assessors (As) #
As2 <- data.frame(ASDR2,V2_1,V2_2,V2_3)
colnames(As2)[1:4]=c("ASDR2","Bias2_beta", "Bias2_lambda", "Bias2_sigma")

eval(parse(text=paste("As2_", data, " <- As2",sep="")))
}                                              # Loop 2 ends

# Compiling #
Stats_f                                                                        <-
rbind(A1=Stats_A1,A2=Stats_A2,A3=Stats_A3,A4=Stats_A4,B1=Stats_B1,B2=Stats_B2,B3=Stats_B3,B4=
Stats_B4,

C1=Stats_C1,C2=Stats_C2,C3=Stats_C3,C4=Stats_C4,D1=Stats_D1,D2=Stats_D2,D3=Stats_D3,D4=Stats_
D4)

Stats2_f                                                                        <-
rbind(A1=Stats2_A1,A2=Stats2_A2,A3=Stats2_A3,A4=Stats2_A4,B1=Stats2_B1,B2=Stats2_B2,B3=Stats2_
B3,B4=Stats2_B4,

C1=Stats2_C1,C2=Stats2_C2,C3=Stats2_C3,C4=Stats2_C4,D1=Stats2_D1,D2=Stats2_D2,D3=Stats2_D3,D
4=Stats2_D4)

As_f <- rbind(A1=As_A1,A2=As_A2,A3=As_A3,A4=As_A4,B1=As_B1,B2=As_B2,B3=As_B3,B4=As_B4,
       C1=As_C1,C2=As_C2,C3=As_C3,C4=As_C4,D1=As_D1,D2=As_D2,D3=As_D3,D4=As_D4)

As2_f                                                                          <-
rbind(A1=As2_A1,A2=As2_A2,A3=As2_A3,A4=As2_A4,B1=As2_B1,B2=As2_B2,B3=As2_B3,B4=As2_
B4,

C1=As2_C1,C2=As2_C2,C3=As2_C3,C4=As2_C4,D1=As2_D1,D2=As2_D2,D3=As2_D3,D4=As2_D4)

# Exporting Results (".csv" format) #
write.csv(Stats_f[i],"Stats_f.csv")
write.csv(Stats2_f[i],"Stats2_f.csv")
write.csv(As_f[i],"As_f.csv")
write.csv(As2_f[i],"As2_f.csv")

eval(parse(text=paste("Stats_f_", i , " <- Stats_f",sep="")))
eval(parse(text=paste("Stats2_f_", i , " <- Stats2_f",sep="")))
eval(parse(text=paste("As_f_", i , " <- As_f",sep="")))
eval(parse(text=paste("As2_f_", i , " <- As2_f",sep="")))
}                                              # Loop 1 ends
```