Technical note: Bayesian calibration of dynamic ruminant nutrition models

K. F. Reed,* G. B. Arhonditsis,† J. France,‡ and E. Kebreab*[1]
*Department of Animal Science, University of California, Davis 95616
†Department of Physical and Environmental Sciences, University of Toronto, Toronto, Ontario, MIC 1A4, Canada
‡Centre for Nutrition Modelling, Department of Animal Biosciences, University of Guelph, Guelph, Ontario N1G 2W1, Canada

ABSTRACT

Mechanistic models of ruminant digestion and metabolism have advanced our understanding of the processes underlying ruminant animal physiology. Deterministic modeling practices ignore the inherent variation within and among individual animals and thus have no way to assess how sources of error influence model outputs. We introduce Bayesian calibration of mathematical models to address the need for robust mechanistic modeling tools that can accommodate error analysis by remaining within the bounds of data-based parameter estimation. For the purpose of prediction, the Bayesian approach generates a posterior predictive distribution that represents the current estimate of the value of the response variable, taking into account both the uncertainty about the parameters and model residual variability. Predictions are expressed as probability distributions, thereby conveying significantly more information than point estimates in regard to uncertainty. Our study illustrates some of the technical advantages of Bayesian calibration and discusses the future perspectives in the context of animal nutrition modeling.
Key words: Bayesian methods, ruminant, mechanistic modeling

Technical Note

Mathematical models of ruminant nutrition are used to formulate diets and integrate knowledge of ruminant digestion and metabolism. Models can be generally classified as empirical or mechanistic, each with different strengths and weaknesses. Empirical models are fitted to "training" data sets, and their application is constrained by the information in the development data. Mechanistic models of ruminant nutrition are causal in nature and can be theoretically applied to draw predictions outside their calibration domain. However, they are deterministic by design and their applications

usually lack parameter variance or prediction error estimates (Baldwin, 1995; Reed et al., 2015). In cases where parameter or prediction errors are estimated through bootstrapping or other fitting techniques, they are usually estimated for predetermined components of the model and do not effectively capture the covariance among all model parameters. Many national nutrient requirement models, such as the NRC (2001), are factorial combinations of stochastic, empirical models that incorporate error assessment but are limited by their static, factorial nature. The factorial approach of sequential combinations of empirically fit models postulates independence among the subsystems. However, in the ruminant animal, the processes of digestion and metabolism are tightly interlinked. In contrast, fitting a whole-animal mechanistic model relaxes the assumption of independence as the parameters determining the behavior of each subsystem are specified simultaneously. Mechanistic models of ruminant nutrition such as those of Baldwin et al. (1987), Dijkstra et al. (1992), and Kebreab et al. (2002) depict our most accurate mathematical representation of the causal relationships operating at the subsystem level to represent the animal physiology dynamically given our current state of knowledge. However, the fitting methods typically used do not explicitly accommodate the error associated with the data nor do they allow for estimation of the appropriateness of model structure.

Models are not a true representation of complex biological systems but rather a depiction of our best understanding of the dominant processes within that system (Oreskes et al., 1994). Recognizing that there are forces acting on the system that are unaccounted for by the model and others that may be described incorrectly, one must explicitly recognize that model predictions will often be inaccurate. For example, the effect of particle size on rate of passage is a force that is not accounted for in this model, which will cause some degree of prediction error. Thus, it is the obligation of the modeler to communicate to the users the level of confidence with which the model can be expected to align with or deviate from reality. As knowledge of ruminant digestion and metabolism increases, the foundation of mechanistic models is strengthened and so

their ability to respond to complex questions that face ruminant production systems today is also enhanced. However, if mechanistic models are to be used as decision-making tools, model errors need to be estimated well to highlight the parts of their structure that are not well-defined and to provide an assessment of the risks associated with the different production system practices.

Bayesian mechanistic models are popular in fields such as ecology and biogeochemistry (Arhonditsis et al., 2008; Ramin and Arhonditsis, 2013) and have many advantages that will benefit the field of ruminant nutrition. Applying Bayesian calibration techniques to dynamic ruminant nutrition models combines the advantageous features of both mechanistic and statistical approaches. Bayesian calibration frameworks can characterize multi-level structures (Zhang and Arhonditsis, 2009): a particularly useful attribute when modeling animal physiology, where individual metabolic processes connect to describe the productivity patterns of individual animals. Another advantage of Bayesian inference techniques is that they allow capitalizing on existing knowledge of the relative plausibility of model parameter values through the formulation of prior distributions (Arhonditsis et al., 2008). The amount of knowledge or confidence about the values of a given parameter determines the degree of information provided by the corresponding prior distribution. Designation of prior distributions constrains the solution space, which in mechanistic models can be prohibitively large and create barriers to model training exercises. Even a small amount of information about a parameter, for example, designation of a distribution that is limited to the positive real line, can prevent the search algorithm from going out of bounds and facilitate convergence to the best solution. Model practitioners have disparate views regarding the use of prior information alongside calibration data. One school of thought encourages the use of noninformative priors, which allows the data to, almost exclusively, determine the posterior parameter estimates. Indeed, in the asymptotic case where the number of observations approaches infinity, Bayesian and frequentist estimation are practically identical if noninformative conjugate prior distributions are used. On the other hand, when sample sizes are small (relative to the number of parameters being estimated), it may not be prudent to solely rely on the available data in guiding the search for defensible model solutions. In this case, the inclusion of prior information on model parameters can be beneficial in that it characterizes the parameter space with respect to its plausibility and therefore effectively reduces the discrepancy between model inputs and outputs (Gelman et al., 2014). Further, when parameter distributions are updated through

model calibration, posterior distributions can then be used as prior distributions for the next calibration when new data become available. Viewing model calibration as an inverse problem, the "prior–likelihood–posterior" update cycles more effectively lead to model solutions that can realistically reflect the internal structure of the modeled system and avoid getting "good results for the wrong reasons" (Zhang and Arhonditsis, 2008).

The Bayesian approach to mechanistic modeling can be adapted to model daily or sub-daily (i.e., hourly or any other period less than a day) time steps in steady-state or non-steady-state dynamics depending on the quality and availability of data. The appropriateness of assuming steady state in simulations of dynamic systems has been questioned (Flynn, 2006). A steady-state model will, at best, achieve an average approximation of the ruminant animal if it is fed several equally proportioned meals a day, which is contrary to common feeding practices of less than 3 meals per day. More realistic feeding patterns can be accommodated in a Bayesian setting through use of dynamic forcing functions, making investigations into the effects of sub-daily fluctuations in rumen fermentation or rate of passage, for example, particularly accessible.

To illustrate this modeling technique, a 3-pool model of rumen N digestion and passage to the duodenum was developed. A description of model equations is given in Table 1. Inputs to the system through N intake are represented by I_X, where X is either the soluble or potentially degradable N fraction in the feed; flows are represented by F_{A-B} where the flow is movement of N from pool A to pool B. State variables or pools are represented by Q and degradation of N from the potentially degradable N pool (Q_{DegN}) to the soluble N pool (Q_{SolN}) is represented as a mass-action function mediated by the degradation constant k_d. Similarly, flow of N from the rumen to the duodenum is also represented as a mass-action function through the rate of passage constant, k_p. The rate of passage (k_p) was further modeled as a function of intake where intake is a function of time:

$$k_p = k_{p0} + \beta_{k_p} \times \left[\frac{I_{DM}(t)}{BW} \times 100 \right], \qquad [1]$$

where β is the coefficient representing the linear relationship between intake and rate of passage. The intake function, $I_{DM}(t)$, was modeled as an interpolation of feed intake over time such that the area under the curve was equal to total DMI. Uptake of N from Q_{SolN} to the microbial N pool (Q_{MicN}) is modeled through a Michaelis-Menten function, similar to that described by Dijkstra et al. (1992), where the maximum rate of

uptake ($V_{\max}$) is a function of the digestible organic matter content of the feed (OMd):

$$V_{\max} = \beta_{V_{\max}} \times \text{OMd}. \qquad [2]$$

Prior distributions for model parameters, $\boldsymbol{\theta} = \left[k_{\mathrm{d}}, k_{\mathrm{m}}, V_{\max}, \beta_{V_{\max}}, k_{\mathrm{p}}, \beta_{k_{\mathrm{p}}} \right]$, were designated as lognormal distributions in which the variance was specified such that the minimum and maximum of a 95% confidence interval were (0.5μ) and 2μ, where μ is an estimate of the parameter obtained from the literature (Table 2). The lognormal distribution was chosen because it excludes biologically impossible negative numbers and the shape and scale parameters are interpreted within the context of the system being modeled. Alternative formulations to the weakly informative prior distributions implemented here include noninformative (or flat) priors such as a normal $(0, 100^2)$ or uniform $(-100, 100)$ distribution. However, given the moderate number of parameters in the model, the reasonable amount of calibration data, and the substantial literature-based biological information about the parameters, we opted for weakly informative priors as the best choice for the present parameter estimation exercise. The calibration data consisted of 17 observations on one animal of Q_{MicN} and the total flow (F_{Tot}) of N out of the rumen to the duodenum taken over the course of a 3-d trial (Robinson, 1983, and personal communication from the author). Total diversion of the digesta at the duodenum was used to measure flow and no record of milk production was taken as it was not the focus of that trial. The diet consisted of 65% mixed alfalfa and grass hay, 26% cracked corn, and 8% soybean meal on a DM basis fed in one meal of 6.55 kg/d. Two Gaussian formulations were considered in which the model was assumed to be either a perfect simulator of the system with the only

error attributed to measurement of the data (M_0; Equation [3]) or an imperfect simulator of the system with error due to both model structure and measurement error (M_1; Equation [4]) are given below:

$$\boldsymbol{y}_{ij} = f\left(\boldsymbol{\theta}, \boldsymbol{y}_0, \boldsymbol{I}_i\right) + \varepsilon_{ij}, \qquad [3]$$

$$\boldsymbol{y}_{ij} = f\left(\boldsymbol{\theta}, \boldsymbol{y}_0, \boldsymbol{I}_i\right) + \varphi_j + \varepsilon_{ij}. \qquad [4]$$

In Equations [3] and [4], $\boldsymbol{y}_{ij}$ is the vector of 17 observations over time for $i = 1, \ldots, 17$ and $j = [Q_{\mathrm{MicN}}, F_{\mathrm{Tot}}]$; $f(\boldsymbol{\theta}, \boldsymbol{y}_0, \boldsymbol{I}_i)$ is the solution of the differential equation system at time point i, given the parameters $\boldsymbol{\theta}$, initial values $\boldsymbol{y}_0$, and model inputs $\boldsymbol{I}$; φ_j is the structural error due to the model; and ε_{ij} is the random error due to measurement. A noninformative Gamma distribution was designated for the error precision parameter φ_j^{-2} because it is constrained to the positive real line and is the conditional conjugate prior when a Gaussian likelihood is assumed. The error due to measurement (ε_{ij}) was specified as 15% of the measured value, $\boldsymbol{y}_{ij}$; the latter specification was used in recent Bayesian calibration exercises in the literature (Arhonditsis et al., 2008; Zhang and Arhonditsis, 2008), whereas the effect of the assumptions made about the measurement error can always be examined through sensitivity analysis (Zhang and Arhonditsis, 2009).

The model was fit in WinBugs software using the WBDiff package for differential equations (Spiegelhalter et al., 2003). Convergence of the Markov chain Monte Carlo chains was checked using the Brooks-Gelman-Rubin scale-reduction factor (Brooks and Gelman, 1998). This diagnostic is based on analyzing multiple simulated Markov chain Monte Carlo chains by comparing the variances within each chain and the variance between chains. Large deviation between these

Table 1. Model parameter descriptions and mathematical representations

Pool/flux	Description	Mathematical description[1]
Q_{SolN}	Rumen soluble N (fraction A)	$I_{\mathrm{Sol}} + F_{\mathrm{Deg-Sol}} - F_{\mathrm{Sol-Mic}} - F_{\mathrm{Sol-Duod}}$
Q_{DegN}	Rumen potentially degradable N (fraction B)	$I_{\mathrm{Deg}} - F_{\mathrm{Deg-Sol}} - F_{\mathrm{Deg-Duod}}$
Q_{MicN}	Microbial N	$F_{\mathrm{Sol-Mic}} - F_{\mathrm{Mic-Duod}}$
I_{Sol}	Input to rumen soluble N	Soluble N fraction $\times$ N intake
I_{Deg}	Input to rumen degradable N	Degradable N fraction $\times$ N intake
$F_{\mathrm{Deg-Sol}}$	Degradation of DegN to SolN	$k_{\mathrm{d}} \times Q_{\mathrm{SolN}}$
$F_{\mathrm{Sol-Mic}}$	Uptake of SolN into MicN	$\dfrac{V_{\max}}{1 + k_{\mathrm{m}} / C_R}$
$F_{\mathrm{Sol-Duod}}$	Passage of SolN to DuodN	$k_{\mathrm{p}} \times Q_{\mathrm{SolN}}$
$F_{\mathrm{Deg-Duod}}$	Passage of DegN to DuodN	$k_{\mathrm{p}} \times Q_{\mathrm{DegN}}$
$F_{\mathrm{Mic-Duod}}$	Passage of MicN to DuodN	$k_{\mathrm{p}} \times Q_{\mathrm{MicN}}$

[1]k_{d} (/h) is the rate of degradation of degradable N into soluble N; $V_{\max}$ (g of N/h) is the maximum rate of incorporation of soluble N into microbial N; k_{m} (g of N/L) is the concentration at half maximal rate of uptake of soluble N into microbial N; C_R (g of N/L) is the concentration of soluble N in the rumen; and k_{p} (/h) is the rate of passage of rumen contents into the duodenum.

2 variances indicates nonconvergence, whereas values close to 1 suggest that each of the multiple chains has stabilized, and they are likely to have reached the target distribution. The latter condition was met for each of the model parameters considered. The accuracy of the posterior parameter values was inspected by ensuring that the Monte Carlo error (an estimate of the difference between the mean of the sampled values and the true posterior mean) for all parameters was less than 15% of the sample standard deviation (Geweke, 1992). For more information on Bayesian modeling in the Bugs platform, see Lunn et al. (2012) and Ntzoufras (2011).

Inference about parameter posterior distributions and the amount of information the data contain about each parameter can be drawn by comparing parameter prior and posterior distributions. This can be done qualitatively through visual assessment of prior and posterior plots (Figure 1) and comparisons of distribution shape and scale parameters (Table 2). The delta index ($\delta_{\theta i}$, Equation [5]) is a quantitative metric that estimates the change in parameter distribution shape and can be conveniently expressed as a percentage of the maximum value of the metric (Endres and Schindelin, 2003; Ramin and Arhonditsis, 2013):

$$\delta_{\theta i} = \sqrt{\int \left[\begin{array}{l} \pi(\theta_i) \log \dfrac{2\pi(\theta_i)}{\pi(\theta_i) + \pi(\theta_i \mid D)} + \\ \pi(\theta_i \mid D) \log \dfrac{2\pi(\theta_i \mid D)}{\pi(\theta_i) + \pi(\theta_i \mid D)} \end{array} \right] d\theta. } \quad [5]$$

Here, $\pi(\theta_i)$ and $\pi(\theta_i \mid D)$ are the prior and posterior distributions of parameter θ_i, respectively. In the Bayesian context, this metric can be interpreted as the square root of twice the information gained about θ_i upon observation of the data, D. The delta index has a minimum value of 0 when the prior and posterior are the same, and a maximum value of $\sqrt{2\log_{10} 2}$ when the 2 distributions do not overlap. The delta indices of model parameters, expressed as a proportion of the maximum value, are given in Table 2.

Model fit can be assessed through the posterior predictive P-value (Meng, 1994; Gelman et al., 1996), the deviance information criterion (Speigelhalter et al., 2002), and the Bayes factor (Kass and Raftery, 1995). The posterior predictive P-value was developed as a Bayesian measure analogous to the frequentist P-value. However, the traditional P-value gives the probability of observing the data under the null hypothesis whereas the posterior predictive P-value is the probability that data replicated from the proposed model is more extreme than the observed data. Thus, in the Bayesian setting, the null hypothesis that the replicated data do not systematically differ from the observed data is rejected when the posterior predictive P-value approaches probabilities of 0 and 1, and is accepted if the P-value is close to 0.5 (Arhonditsis et al., 2008). Two measures of discrepancy, the χ^2 test statistic and the log-likelihood, were used to assess the posterior predictive P-value of the model (Gelman et al., 1996). The P-value can be expressed visually by plotting the discrepancy of the data from the model against the discrepancy of the posterior predictions from the model (Figure 2a,c). An alternative visualization is to produce the histogram of the likelihood values of the posterior predicted replicates relative to the model with the mean likelihood of the data (given the model) plotted as a vertical line (Figure 2b,d). It can also be expressed as a numerical value representing the number of cases in which the likelihood of the posterior predictions is greater than the mean likelihood of the data (Figure 2; Gelman et al., 1996).

Table 2. Comparison of prior and posterior densities[1] from model 1 (Equation [4])

Parameter[2]	Central tendency		SD		
	Prior (median)	Posterior (mean)	Prior	Posterior	Delta index
k_d	0.14	0.137	0.984	0.0363	0.664
$\beta_{V_{max}}$	21.03	13.9	122.4	2.87	0.303
k_m	3	4.39	18.4	1.13	0.505
β_{k_p}	0.05	0.0626	0.284	0.0185	0.169
k_{p0}	0.09	0.0890	0.530	9.95E-3	0.397

[1]Prior densities are user defined likelihood distributions of parameter values based on knowledge of the system. Posterior values are estimates of parameter likelihoods after model calibration. The delta index is a standardized measure of how much the parameter likelihood changed through calibration with 0 indicating no change and 1 indicating maximum possible change.

[2]k_d (/h) is the rate of degradation of degradable N into soluble N; $\beta_{V_{max}}$ [g of N/(h·g of dOM)] is the rate of increase in maximum rate of incorporation of soluble N into microbial N per unit of digestible OM; k_m (g of N/L) is the concentration at half maximal rate of uptake of soluble N into microbial N; β_{k_p} [/(h·g of DMI)]; k_{p0} (/h) is the initial rate of passage of rumen contents into the duodenum.
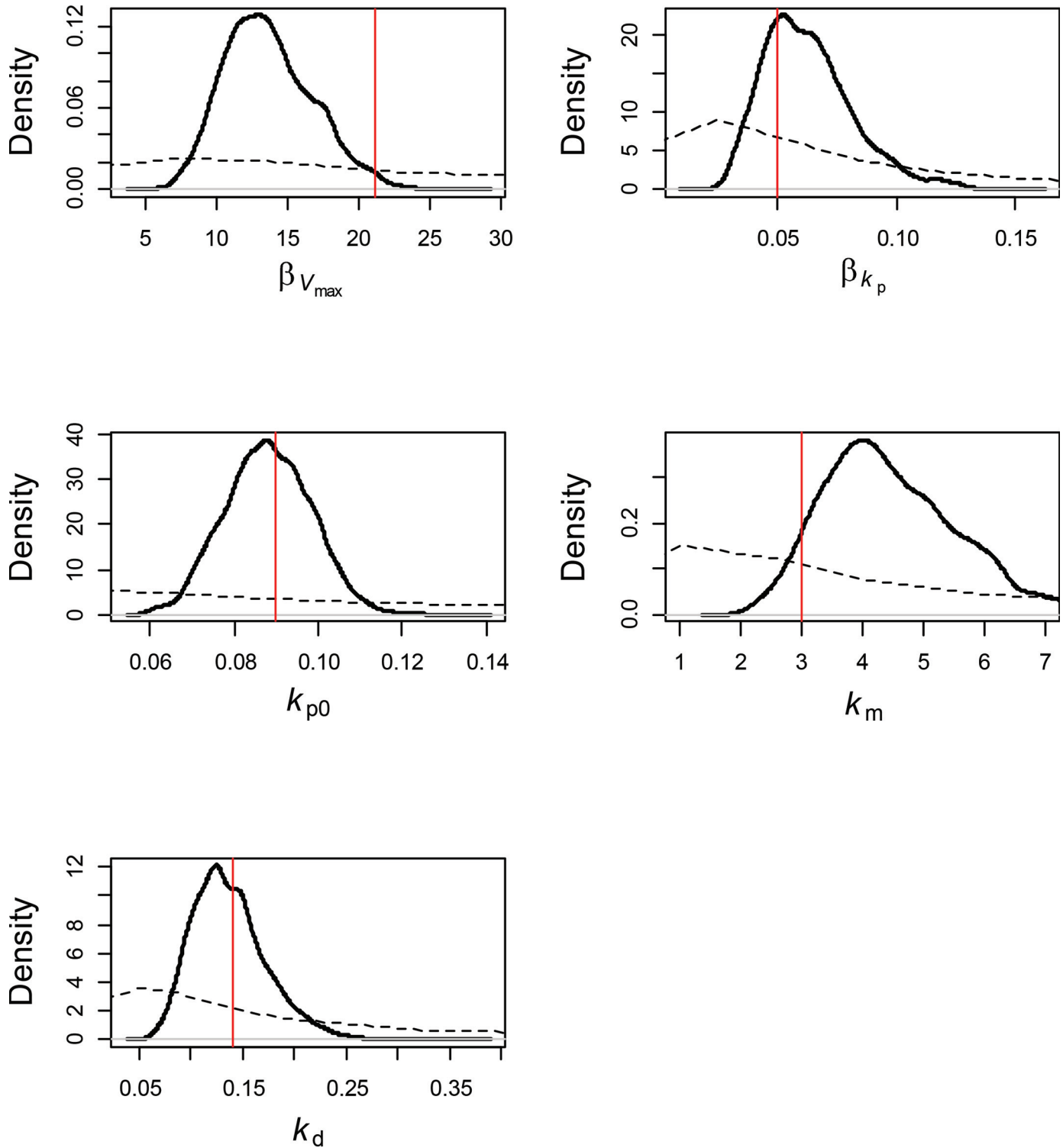
Figure 1. Parameter prior (dashed lines) and posterior (solid black line) distributions with the prior medians shown in the vertical line. $\beta_{V_{max}}$ [g of N/(h·g of dOM)] is the rate of increase in maximum rate of incorporation of soluble N into microbial N per unit of digestible OM; k_{p0} (/h) is the initial rate of passage of rumen contents into the duodenum; k_d (/h) is the rate of degradation of degradable N into soluble N; β_{k_p} (/h × kg of DMI) is the coefficient representing the linear relationship between intake and rate of passage; k_m (g of N/L) is the concentration at half maximal rate of uptake of soluble N into microbial cells. Color version available online.
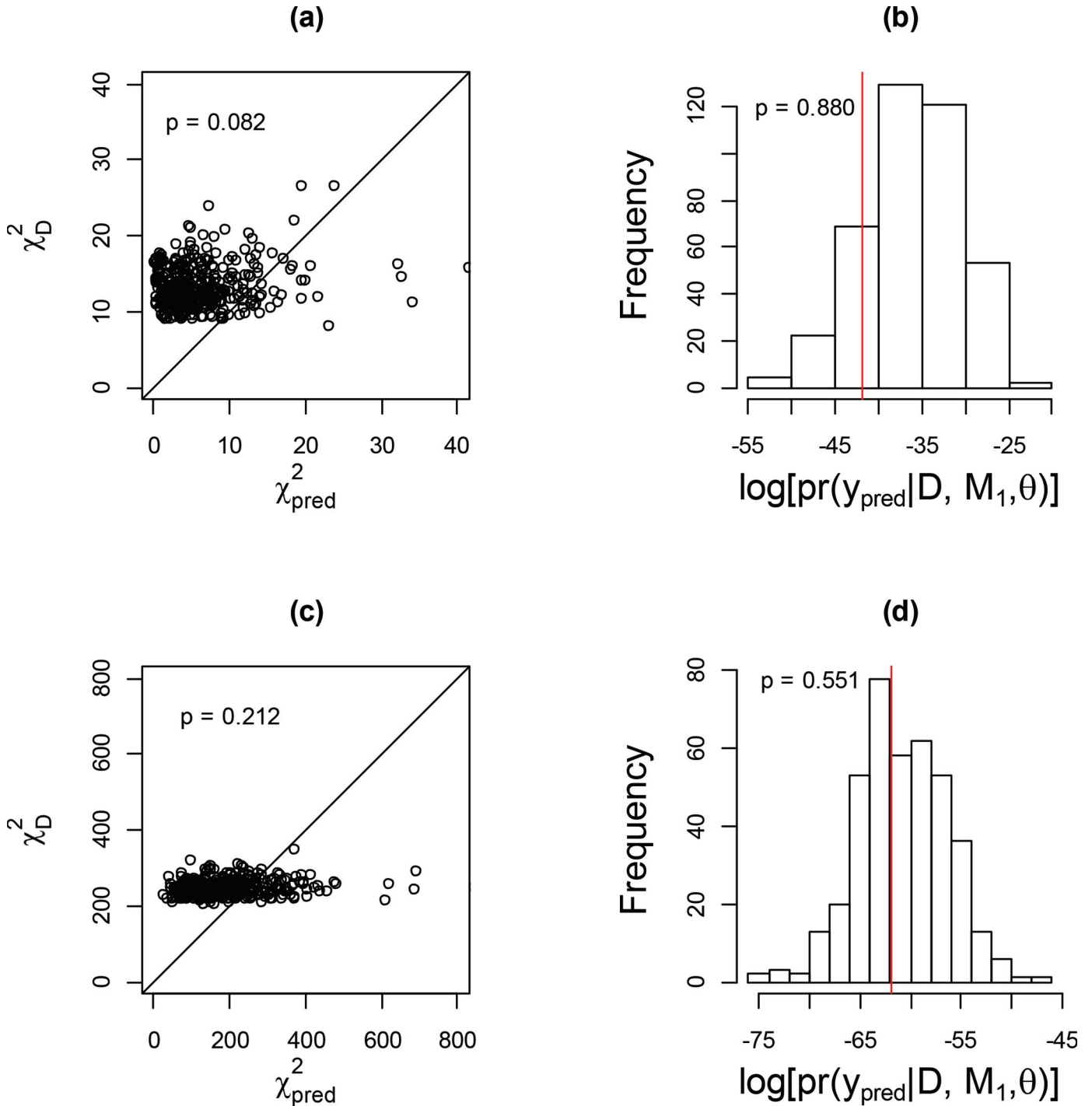
Figure 2. Posterior predictive *P*-values as assessed by the χ^2 test statistic (a, c) and the log-likelihood (b, d) of the data (D) and model predicted replicates (y_{pred}) for the microbial nitrogen pool (a, b) and total flux of nitrogen to the duodenum (c, d). In the log-likelihood histograms, the mean data log-likelihood given the model at hand is shown with a vertical line. M_1 designates the model formulation used that included measurement and structural model error as indicated equation 4; θ is the vector of prior distributions. Color version available online.

The latter 2 measures are means for comparing model fitness when more than one model is investigated. The deviance information criterion is the Bayesian counterpart of the Akaike information criterion in

that both consider model performance while penalizing complexity and thus favoring selection of the most parsimonious models (Speigelhalter et al., 2002). The Bayes factor (B_{10}) is the posterior odds of one model

(M_1) over the other (M_0), which is estimated from the model likelihood of the data and the (often equal) prior probabilities assigned to each model of interest using Bayes theorem:

$$B_{10} = \frac{\text{pr}(D \mid M_1)\,\text{pr}(M_1)}{\text{pr}(D \mid M_0)\,\text{pr}(M_0)}. \qquad [6]$$

Here, M_0 represents the null hypothesis (H_0) that the model is a perfect simulator of the system (Equation [3]), and M_1 represents the alternative (H_1), that there is a structural error due to the model (Equation [4]). Assuming $\text{pr}(M_1) = \text{pr}(M_0)$, B_{10} can be estimated simply from the ratios of model likelihoods where the posterior model likelihood is integrated over the posterior distribution of model parameters. In this example, incorporation of a structural error term (M_1, Equation [4]) resulted in a Bayes factor > 100 which, according to the framework proposed by Kass and Raftery (1995), is decisive evidence in support of M_1 over M_0.

Plots of model (M_1) predictions and observed values over time (Figure 3) provide a visual assessment of the model fit to the data and the expected pattern of pool fluctuation under a single daily feeding schedule. Most observations fall within the 95% credible intervals (Figure 3) indicating a good model fit, although the derived uncertainty bands are fairly broad and extend to less than zero for F_{Tot}. Whereas the latter posterior pattern primarily stems from the substantial temporal variability of the calibration data set, we note that these credible intervals serve as a proof of concept of the ability of the proposed calibration framework to provide dynamic predictive error estimates. From a statistical standpoint, the negative values for F_{Tot} can be avoided by truncating the Gaussian model likelihood used or by simply postulating that the model error is log-normally distributed. On the other hand, from an inferential viewpoint, the likelihood of negative predictive outputs during the parameter estimation exercise could have implications about the quality of the calibration data, especially when we use multivariate data compiled from different sources, the mechanistic foundation of the mathematical model developed, or both.

To illustrate the utility of Bayesian mechanistic models in predicting future feeding or management scenarios, posterior estimates were used to run simulations and assess model outcomes. Markov chain Monte Carlo samples of parameters were used to construct a posterior multivariate distribution, based on parameter mean values and variance-covariance patterns (Arhonditsis et al., 2008). Two dietary inputs to the model were varied to assess their effect on microbial protein production. Specifically, OMd content and SolN

as a percentage of total N were varied in even increments from 0.55 to 0.85 and 0.18 to 0.5, respectively. A Markov chain Monte Carlo of the microbial N pool was monitored and recorded for daily microbial N production under each dietary setting. The proportion of the Markov chain Monte Carlo chain that exceeds the designated threshold of 90 g/d provides the likelihood that microbial N production exceeds the threshold in the given dietary scenario. The threshold was chosen as a reasonable estimate of microbial N production for the given DMI based on values reported in the literature for higher levels of DMI and the NRC (2001) estimate of microbial N production for the calibration data diet. The results of the simulations are given in Figure 4 and indicate that increasing OMd and SolN both increase the probability that microbial N will exceed the threshold.

Model evaluation results suggest the model fits the data well and the data contain sufficient information to improve estimates for most of the parameters. The delta indices given in Table 2 and density plots in Figure 1 indicate that even the small amount of data used in this study provide enough information to change the density of the parameters, resulting in a larger increase in confidence of parameter estimates. The posterior
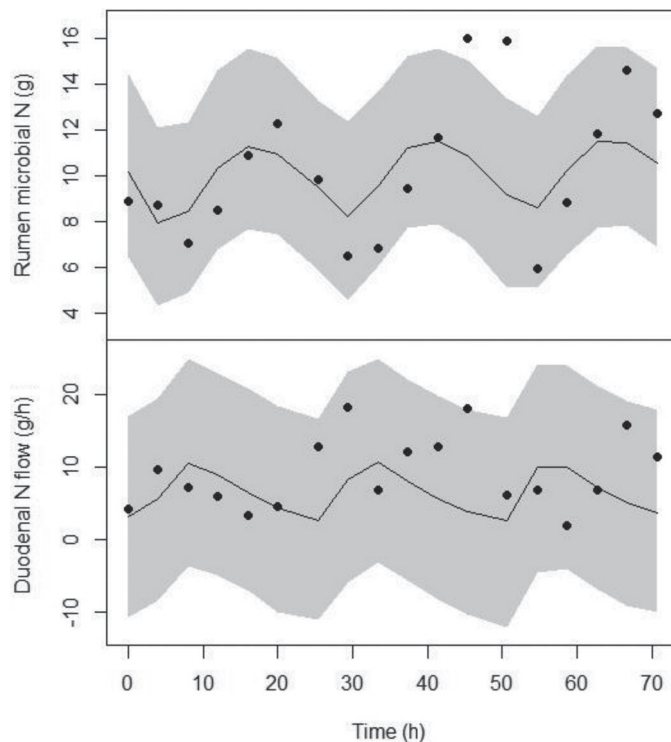

Figure 3. Model solutions (lines) and 95% credible intervals (gray area) plotted against time (h) along with the observed data of the microbial N pool and N flow to the duodenum used to fit the model (symbols).
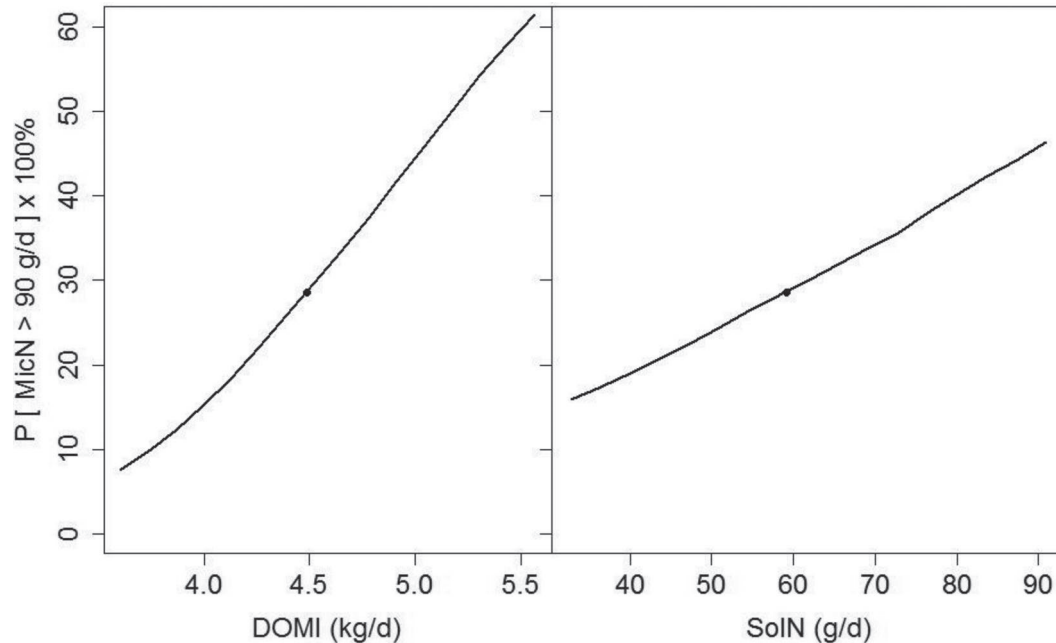
Figure 4. Posterior probability of microbial nitrogen production (MicN, g/d) exceeding 90 g/d under varying digestible organic matter intake (DOMI, kg/d) and soluble nitrogen intake (SolN, g/d). Points represent values of the observed diet.

predictive *P*-values shown in Figure 2 indicate that model predictions do not systematically differ from the observed values for Q_{MicN} or F_{Tot}. However, grouping the observations over time to get one estimate of model fit for each type of data could mask a systematic error or lack of fit at a specific time point. To avoid this problem, a posterior predictive *P*-value for each time point and each type of data should be estimated. The large (>100) Bayes factor when comparing M_0 and M_1 indicates that accounting for model structural error greatly improves model fit. This error structure can be expanded to allow the error to vary over time or to accommodate other sources of error such as individual animal or study variance. The Bayes factor can then be used to select the best model for the data. Finally, model predictions plotted over time with observations (Figure 3) support a good model fit and follow the expected pattern of digesta flow for a single-daily feeding regimen.

The posterior simulations illustrated in Figure 4 communicate the degree of confidence that a certain threshold will be exceeded. From a user perspective, this uncertainty is a valuable piece of information that could prevent loss of faith in model utility when predictions are not realized. For example, instead of basing our decisions on point (average) estimates of milk production without any insights into the associated uncertainty, the Bayesian calibration framework can predict the probability with which an examined diet will exceed

a given level of milk production (e.g., there is an 80% probability that milk production will be higher than 35 kg/cow per d, if we adopt a particular diet). In essence, the Bayesian proposition aims to shift the focus of the decision making process from an *average prediction* to a *probability mass* that delineates the likelihood of success (or failure) of a tested production practice. It is the responsibility of animal scientists to provide users with the most credible information available to guide their management decisions, but also to be transparent about the level of precision and accuracy associated with model predictions.

The methods described provide a framework for fitting a simple 3-pool model that accounts for error due to measurement and model structural error for each source of data. This model can be expanded (given data availability) by adding additional pools, errors that vary over time, random effects of animal or study (or both), or modeling parameter variances. The flexibility of the Bayesian framework and the ability to probabilistically incorporate prior knowledge of the system makes this technique well suited to mechanistic models of ruminant digestion and metabolism. In most experiments, data are collected at many time points over the course of a study; however, much of the information is lost when these observations are averaged by day or trial period. Bayesian mechanistic models can be fit to time series allowing all of the data to inform model estimates. Further, the iterative nature of the Bayes' theorem, which

allows sequential updating of parameter distributions, facilitates accumulation of information from the data through multiple calibrations (Arhonditsis et al., 2008). Finally, the posterior uncertainty patterns can be used to guide data collection efforts and experimentation that revolves not only around model endpoints but also individual mechanisms or process rates (Zhang and Arhonditsis, 2008). User-friendly software, advances in computing power, and a growing body of work from other fields make this approach more accessible than ever.

The challenge facing the dairy industry today is to increase production in an economically and environmentally sustainable manner. To meet this challenge, producers and policy makers depend on scientific analysis of increasingly complex systems. A large amount of uncertainty about the principles that direct ruminant digestion and metabolism, including everything from rumen fermentation to genetic variation, exists. As long as this uncertainty remains, it is the modeler's responsibility to provide a cogent analysis of the contribution of individual sources and collectively evaluate the degree of confidence to model predictions. Conventional deterministic mechanistic models do not provide error estimates and thus fail to provide an assessment of the risk associated with decisions based on model results. Bayesian mechanistic models capture the inherent variability of the biological system under study and provide an assessment of the error associated with complex model results. Decision makers such as producers can benefit from knowledge of associated error while predicting outcomes such as milk production.

ACKNOWLEDGMENTS

REFERENCES

Arhonditsis, G. B., D. Papantou, W. Zhang, G. Perhar, E. Massos, and M. Shi. 2008. Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management. J. Mar. Syst. 73:8–30.

Baldwin, R. L. 1995. Modeling Ruminant Digestion and Metabolism. Chapman and Hall, New York, NY.

Baldwin, R. L., J. H. Thornley, and D. E. Beever. 1987. Metabolism of the lactating cow II: Digestive elements of a mechanistic model. J. Dairy Res. 54:107–131.

Brooks, S. P., and A. Gelman. 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7:434–455.

Dijkstra, J., H. D. St. C. Neal, D. E. Beever, and J. France. 1992. Simulation of nutrient digestion, absorption, and outflow in the rumen: Model description. J. Nutr. 122:2239–2256.

Endres, D. M., and J. E. Schindelin. 2003. A new metric for probability distributions. IEEE Trans. Inf. Theory 49:1858–1860.

Flynn, K. J. 2006. Reply to Horizons article 'Plankton functional type modelling: running before we can walk' Anderson (2005): II. Putting trophic functionality into plankton functional types. J. Plankton Res. 28:873–875.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2014. Bayesian Data Analysis. CRC Press, New York, NY.

Gelman, A., X. Meng, and H. Stern. 1996. Posterior predictive assessment of model fitness via realized discrepancies. Stat. Sin. 6:733–807.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In Bayesian Statistics 4. J. M. Bernado, J. O. Berger, A. P. Dawid, and A. F. M. Smith, ed. Clarendon Press, Oxford, UK.

Kass, R. E., and A. E. Raftery. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.

Kebreab, E., J. France, J. A. N. Mills, R. Allison, and J. Dijkstra. 2002. A dynamic model of N metabolism in the lactating dairy cow and an assessment of impact of N excretion on the environment. J. Anim. Sci. 80:248–259.

Lunn, D., C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter. 2012. The BUGS Book: A Practical Introduction to Bayesian Analysis. CRC Press, Boca Raton, FL.

Meng, X. 1994. Posterior predictive p-values. Ann. Stat. 22:1142–1160.

NRC. 2001. Nutrient Requirements of Dairy Cattle. 7th rev. ed. Natl. Acad. Sci., Washington, DC.

Ntzoufras, I. 2011. Bayesian modeling using WinBUGS. John Wiley & Sons, Hoboken, NJ.

Oreskes, N., K. Schrader Frechette, and K. Belitz. 1994. Verification, validation, confirmation of numerical models in the earth sciences. Science 263:641–646.

Ramin, M., and G. Arhonditsis. 2013. Bayesian calibration of mathematical models: Optimization of model structure and examination of the role of process error covariance. Ecol. Inform. 18:107–116.

Reed, K. F., D. P. Casper, J. France, and E. Kebreab. 2015. Prediction of nitrogen efficiency in dairy cattle: A review. CABI. http://dx.doi.org/10.1079/PAVSNNR201510001.

Robinson, P. H. 1983. Development and initial testing of an in vivo system to estimate rumen and whole tract digestion in lactating dairy cows. Doctoral Diss. Cornell University, Ithaca, NY.

Spiegelhalter, D., A. Thomas, N. Best, and D. Lunn. 2003. WinBUGS User Manual, Version 1.4. Accessed Mar. 1, 2015. http://www.mrc-bsu.cam.ac.uk/bugs.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. J. Royal Stat. Soc. B Stat. Meth. 64:583–639.

Zhang, W., and G. B. Arhonditsis. 2008. Predicting the frequency of water quality standard violations using Bayesian calibration of eutrophication models. J. Great Lakes Res. 34:698–720.

Zhang, W., and G. B. Arhonditsis. 2009. A Bayesian hierarchical framework for calibrating aquatic biogeochemical models. Ecol. Modell. 220:2142–2161.