J. Dairy Sci. 99:1–9
http://dx.doi.org/10.3168/jds.2015-10708

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 foun-

Received December 2, 2015.
Accepted April 6, 2016.
[1]Corresponding author: ekebreab@ucdavis.edu

dation 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 $F_{\mathrm{A-B}}$ 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:

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

where β is the coefficient representing the linear relationship between intake and rate of passage. The intake function, $I_{\mathrm{DM}}(t)$, 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