Assessing Uncertainty in Mechanistic Models

Edwin J. Green, David W. MacFarlane, and Harry T. Valentine

Abstract.—Concern over potential global change has led to increased interest in the use of mechanistic models for predicting forest growth. The rationale for this interest is that empirical models may be of limited usefulness if environmental conditions change. Intuitively, we expect that mechanistic models, grounded as far as possible in an understanding of the biology of tree growth, may be more useful in an altered environment. Unfortunately, such models often produce only point estimates, with no associated credible or confidence intervals. The Bayesian synthesis (BSYN) method provides a solution to this dilemma. We present a summary of the BSYN, and the results of an application of BSYN to PIPESTEM, a mechanistic model of forest growth calibrated for loblolly pine.

There has recently been a great deal of interest in characterizing the uncertainty associated with mechanistic models. Such models are not constructed from data in the way usual stochastic models are, and hence the common measures of uncertainty are not applicable. In this paper, we demonstrate the Bayesian synthesis method, first proposed by Raftery *et al.* (1995), for developing interval estimates with associated probability statements. We will use a variant of Raftery *et al.*'s original method to obtain predictions from PIPESTEM, a mechanistic model of forest stand growth.

PIPESTEM

The mechanistic model PIPESTEM describes the carbon balance and growth of an even-aged mono-specific stand on an annual basis (Valentine 1988, 1990; Valentine *et al.* 1997). The version we used was calibrated for loblolly pine. The mathematical model consists of 16 differential equations and additional ancillary functions. The differential equations completely describe the dynamic properties of a single model stand. The state variables that characterize the model stand ordinarily are initialized with measurements from a real stand that has aged 1 or more years since planting. Simultaneous numerical integration of the differential equations provides values of the state variables at any subsequent point in time, and these values serve as predictions for the real stand.

The growth rate of the model stand is defined as the difference between rates of production and loss of dry matter. Dry matter is measured in units of carbon *C*, and the rate of growth per unit land area is measured in units of *C/ha/year*. The rate of production of dry matter is defined as the difference between the rate of production of *C* substrate and the consumption of *C* substrate through maintenance and constructive respiration. These metabolic rates also have dimensions with units of *C/ha/year*.

Dry matter is divided into foliar, feeder-roots, and woody components. The proportion of *C* substrate allocated to the production of dry matter of each component is constrained to keep the morphological dimensions of the model stand in accord with pipe model theory (Shinozaki *et al.* 1964a,b). Losses of dry matter result from the turnover of foliage and feeder-roots, the death and self-pruning of branches associated with crown rise, and the death of trees associated with self-thinning.

In addition to the information about the carbon balance and production and loss of dry matter, PIPESTEM also furnishes total basal area (m^2/ha), average tree height (m), dominant tree height (m), average height to the base of a live crown (m), and tree density ($stems/ha$) on an annual timestep. Estimates or measurement of these four variables (ordinarily obtained from a real stand aged 1 year or more) suffice to initialize the PIPESTEM model.

Valentine *et al.* (1997) modeled the growth of loblolly pine stands at three sites in Virginia and one site in North Carolina. These areas were chosen because they are the locations of sample plots for the Virginia Tech Loblolly Pine Growth and Yield Research Cooperative, and the data from the sample plots could be used to evaluate the predictions from PIPESTEM. Predictions from the model matched observed data well.

In this study, we use a version of PIPESTEM that includes a routine that allows the user to specify the site index (average height of dominant trees at 25 years of age) of the stand under study. The total number of model parameters in this version of PIPESTEM is 24. Users must specify values for all 24 parameters, and then the model produces predictions according to the specified parameter values.

Professor and Graduate Assistant, Rutgers University, New Brunswick, NJ, USA; and Research Forester, USDA Forest Service, Northeastern Research Station, Durham, NH, USA, respectively.

BAYESIAN SYNTHESIS

Raftery *et al.* (1995) developed the Bayesian synthesis (BSYN) method to characterize uncertainty in mechanistic models. The basic approach is to use *all* available information to generate a joint premodel distribution on all model inputs (θ) and model outputs (ϕ). In the spirit of Bayesian analysis, this distribution should reflect as closely as possible the investigator's state of knowledge regarding both θ and ϕ *before* running the model (knowledge gained from the model may not be included or it would eventually be used twice). In fact, as we will see, the premodel distribution will often be a Bayesian posterior distribution. The mechanistic model is then used to translate the joint premodel distribution to a joint postmodel distribution on θ and ϕ.

Mathematically, the model consists of a mapping from θ to $\Phi(\theta)$, where $\Phi(\theta)$ is a submanifold in (ϕ-space, i.e., the model connects particular θ values with particular associated ϕ values. Now, let $p(\theta, \phi)$ represent the joint premodel distribution of θ and ϕ, and let $\pi(\theta, \phi)$ be the joint postmodel distribution. In BSYN, we assume the model is correct. This leads to:

$$\begin{aligned} \pi(\theta,\phi) &\propto p(\theta, (\Phi(\theta)) \text{ if } \phi = \Phi(\theta), \\ &\propto 0 \text{ otherwise.} \end{aligned}$$

The essential idea of BSYN is that the postmodel distribution is proportional to the premodel distribution for any values of θ and ϕ that are connected by the model. Other combinations of θ and ϕ are impossible according to the model, and hence are assigned postmodel probabilities of 0. Inference about ϕ or θ or any function of ϕ and/or θ is based on $\pi(\theta, \phi)$. The joint postmodel distribution can be marginalized to obtain a postmodel distribution of any quantity of interest. Interval estimates with valid posterior probability statements are available from the marginalized posterior distributions. For instance, if we let $\pi(\phi)$ denote the marginal postmodel distribution of ϕ, then the postmodel probability that ϕ is in some interval I is simply

$$\text{Prob}(\phi \in I) = \int_I \pi_\phi(\phi)\, d\phi.$$

Note that BSYN yields postmodel distributions not only of the model outputs (ϕ), but also of the model inputs (θ). The latter may prove helpful in future modeling efforts. For more details, consult Raftery *et al.* (1995) (including discussion and rejoinder).

Unfortunately, the original BSYN method suffered from Borel's paradox (Wolpert 1995). Essentially, this means that the results were not invariant to re-parametrization of the model. Givens and Bravington (1995) demonstrated that this is probably not cause for great alarm, but nevertheless it is bothersome. Raftery, Poole, and Givens (1996) (hereafter referred to as RPG) demonstrated that the root problem is a basic incoherence in the BSYN method, which had heretofore gone unnoticed. The problem may be stated as follows: the user specifies premodel distributions on the model inputs and outputs. In essence, the model is a transformation from θ to ϕ. Hence the model and premodel distributions of the inputs provide an *implied* premodel distribution on the outputs. Of course, the implied distribution will not be identical to the specified premodel distribution on the outputs. Hence the incoherence: two different *prior* distributions for the same quantity.

Fortunately, the Borel paradox may be avoided. As mentioned earlier, the joint premodel distribution represents all that is known about θ and ϕ exclusive of the model. This knowledge must have been gained either subjectively or from data. Hence it is appropriate to consider a general situation where information on θ or ϕ or both is represented by a traditional Bayesian subjective prior distribution and/or a likelihood based on whatever data are available. We will also assume that our premodel information on θ and ϕ is from independent sources, although not strictly necessary, as are our data (this will be discussed further in the Application section). Then the joint premodel distribution may be thought of as composed of four parts: two likelihoods and two priors, i.e.,

$$p(\theta, \phi) \propto q_\theta(\theta)\ L_\theta(\theta)\ q_\phi(\phi)\ L_\phi(\phi),$$

where $p(\theta, \phi)$ is the joint premodel distribution; $q_\theta(\theta)$ and $q_\phi(\phi)$ are premodel prior distributions on θ and ϕ, respectively; and $L_\theta(\theta)$ and $L_\phi(\phi)$ are likelihoods for θ and ϕ, respectively. Since likelihoods are known to be invariant to transformation (e.g., see Schweder and Hjort 1996), the incoherence must arise from the premodel prior distributions.

RPG and Givens and Roback (1997, hereafter referred to as GR) suggested a solution to the Borel paradox: pool the implied premodel distribution with the specified premodel distribution. For example, it is easy to find (or draw an arbitrarily large sample from) $q^p_\phi(\phi)$. This could be done by drawing a large sample for θ from $q_\theta(\theta)$. For each sampled observation, run the model to obtain the associated value of ϕ. The collection of ϕ's would then constitute a sample from $q^i_\theta(\phi)$. Importance sampling (see, for example, Rubin 1988) could then be used to select a sample from $q^p_\phi(\phi)$. Then BSYN could be used to derive $\pi_\phi(\phi)$, the postmodel distribution for ϕ.

Unfortunately, unless the model is 1:1, the postmodel distribution for θ is not immediately available from $\pi_\phi(\phi)$ because numerous points in θ-space may map onto the same point in ϕ-space. Hence, while we might be able to approximate the postmodel probability of some point, say, ϕ_p, it is not readily apparent how this information might

be used to obtain postmodel probabilities for all the θ_j's that might have led to ϕ_i.

Raftery and Poole (1997) and Poole and Raftery (1998), hereafter referred to as RP and PR, respectively, developed the "full pooling" method for pooling the stated and implied priors on both the inputs and outputs, resulting in samples from $\pi_\theta(\theta)$ and $\pi_\phi(\phi)$, the postmodel distributions of θ *and* ϕ, respectively. See Green *et al.* (1998) for more details.

APPLICATION

PIPESTEM is calibrated for loblolly pine. Given initial stand conditions (site quality, average diameter, height, height to the live crown, number of stems per hectare, and stand age), the model projects the growth of the stand over time. Among the outputs produced by the model are basal area (m^2/ha) (B), average stand height (m) (H), and number of stems per hectare (N). We used PIPESTEM to simulate the growth of a model loblolly pine stand on the Upper Coastal Plain of Virginia.

Spacing Trial Data

We used the Virginia Tech Spacing Trial data, maintained by the Virginia Tech Loblolly Pine Growth and Yield Research Cooperative. These data include three replications of a number of spacings. Measurements available include average stand height, basal area, and number of stems per hectare. We arbitrarily assumed our model stand was planted at location 3 (Roanoke Rapids, VA). Data from plots at this location were used to compute $L_\phi(\phi)$, the likelihood for the outputs. We did not specify a likelihood of the inputs, since we had no data to base one on. In essence, this means we assumed a constant likelihood for θ.

Premodel Distribution

Likelihood for Model Outputs

Four of the spacings in the Spacing Trial data yielded identical planting densities of 2,242 *stems/ha*. Hence we chose this as our planting density. At the time of this research, the latest data available were for age 12; therefore, we used age 12 data. The likelihood was assumed to be a three-dimensional normal distribution, and the 12 observations (three replications, four spacings) on basal area per hectare, average stand height, and number of stems per hectare were assumed to be independent realizations from this distribution. The mean basal area, height, and number of stems on these 12 plots were 31.5 m^2/ha, 10.9 m, and 2,101.3 *stems/ha*, respectively.

Prior Distribution on Model Inputs

In Bayesian Synthesis, the premodel prior distributions on the inputs, $q_\theta(\theta)$, should reflect all that is known about their feasible values. Consequently, we undertook the arduous task of researching each of the 24 inputs for PIPESTEM. The results of this study are contained in MacFarlane *et al.* (1998). For brevity, we will summarize the findings here.

We assumed the inputs were premodel independent. We recognized that this was unlikely to be true, but we were unable to specify reasonable prior covariances. Hence we decided to allow the model and available data to induce covariances in the postmodel distributions (in fact, one of the virtues of Bayesian analysis is that such covariances are available at the end of the procedure). Note that assuming prior independence may actually be viewed as a *conservative* procedure. Imagine the prior distribution in two dimensions, and for simplicity, assume they have the same scale. If the two parameters are independent, then the prior distribution is represented graphically by a circle. In contrast, if we specify an *a priori* correlation, then the prior distribution would be an ellipse contained within the circle. Hence the prior independence assumption actually allows one to consider more combinations of the two parameters than would be included if *a prior* dependence structure was specified.

For 19 of the parameters, we were able to specify prior means and variances. We used normal priors for these parameters. For two others, z_w and z_r, (see Valentine *et al.* 1997 or MacFarlane *et al.* 1998 for a complete listing of the PIPESTEM parameters and their interpretations) we were able to specify upper and lower bounds in addition to means and standard deviations. Hence, rather than using unbounded normal densities for these parameters, we chose to use beta densities. The beta density parameters were computed from the bounds, means, and standard deviations using the method described in Johnson and Kotz (1970, p. 44).

For the final three parameters (ρ, μ_1, and μ_2) our premodel prior distributions were uniform. We could find no published information on these three parameters. Hence we made liberal guesses about the minimum and maximum values, and specified the distributions to be flat in between.

Prior Distribution on Model Outputs

We obtained a site index of 23.4 m (base age 25) for location 3 from the Amateis *et al.* (1984) Coastal Plain site index equation. With this site index and the planting

density, we derived prior distributions on three of the model outputs (basal area per hectare, average stand height, and number of stems per hectare) at age 12 with the aid of four common growth and yield models (Bailey *et al.* 1985, Amateis *et al.* 1984, Clutter *et al.* 1984, and Hafley *et al.* 1982). Unfortunately, it is not common for model developers to provide sufficient information to develop variance estimates. Hence we assumed a normal prior distribution on the outputs with a mean equal to the average of the predictions from the four models. For the covariance matrix, we used the maximum likelihood estimate from observations on these variables from plots planted at densities of 2,242 *stems/ha* at locations 1, 2, and 4 in the Spacing Trial data.

Model Stand Specification

In all runs of PIPESTEM, we initialized our stand at 2,242 *stems/ha* on land with a site index of 23.4 *m*. The initial (age 1) average diameter and height were specified to be equal to the averages calculated from the plots planted at 2,242 *stems/ha* at location 3 (1.12 *cm* and 0.41 *m*, respectively). Initial average height to the base of the live crown was arbitrarily specified to be 0.1 m.

Details of Bayesian Synthesis for PIPESTEM

We used the full pooling algorithm of Raftery and Poole to generate a sample of size 5000 from $\pi_\theta(\theta)$. Note that the sample from $\pi_\theta(\theta)$ may be thought of as a sample of model trajectories. Thus, although $\pi_\phi(\phi)$ is a sample from the postmodel distribution of ϕ at age 12, we can easily obtain postmodel distributions for ϕ at *any* age by running the model to the desired age for each combination of inputs in the sample from $\pi_\theta(\theta)$.

RESULTS

Histograms from the postmodel distributions of *B*, *H*, and *N* at ages 12 and 50 are presented in figure 1. With these samples it is possible to construct valid credible intervals. For example, in table 1 we present the mean, and the 2.5 and 97.5 percentiles for the samples in figure 1. The percentiles may be regarded as end points of approximate 95 percent credible intervals.

Several of the histograms and/or smoothed densities suggest bimodal distributions. The skew in the posterior

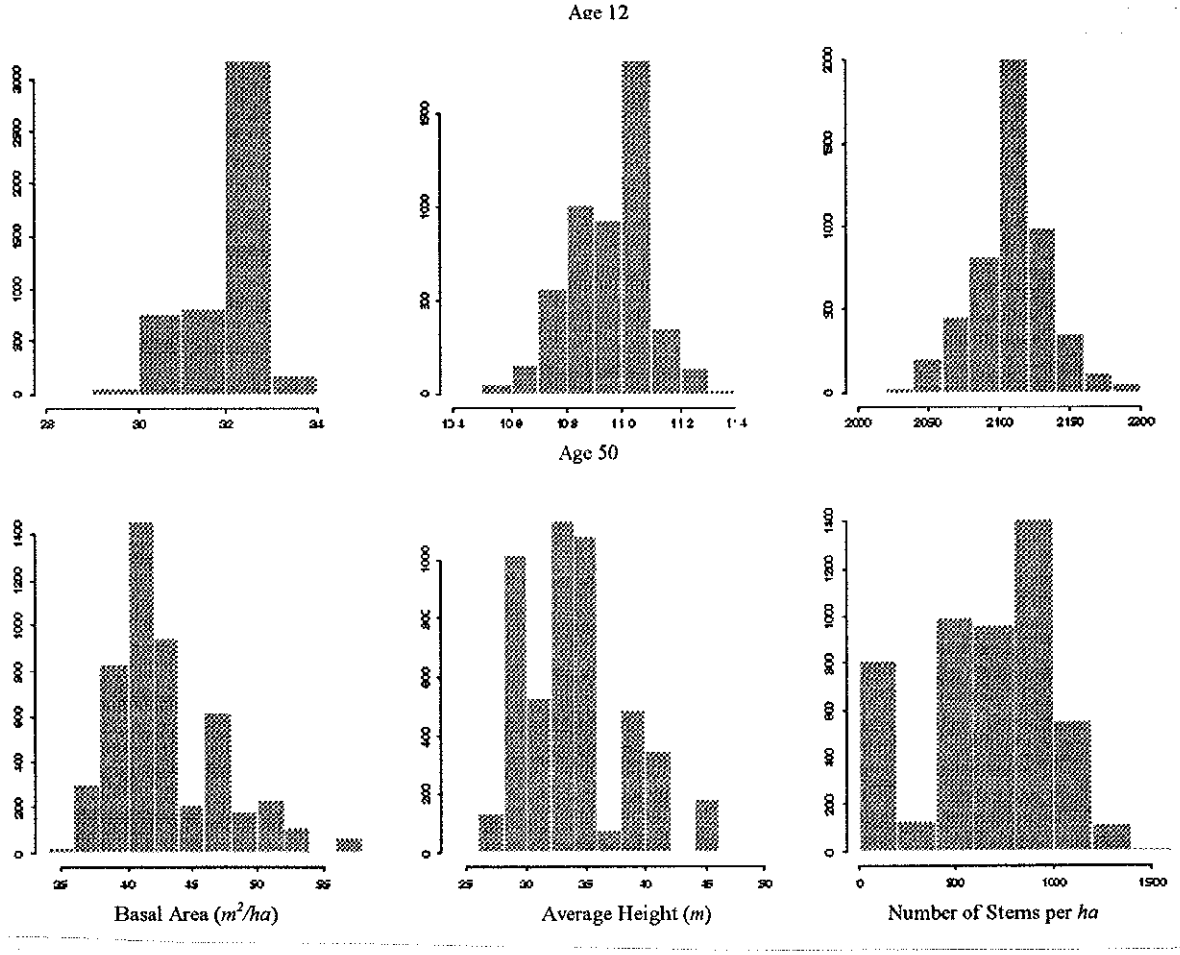

Figure 1.—*Histograms of samples from marginal posterior distribution for basal area (m^2/ha), average height (m), and number of stems per ha, at ages 12 and 50.*

Table 1.—*Means and approximate 95 percent credible intervals from posterior distributions*

Attribute	Mean	Lower limit (2.5 percentile)	Upper limit (97.5 percentile)
B (m^2/ha), age 12	31.9	30.1	33.2
B (m^2/ha), age 50	43.1	36.9	53.8
H (*m*), age 12	11.0	10.7	11.2
H (*m*), age 50	34.1	27.8	45.2
N (*stems/ha*), age 12	2,108	2,057	2,174
N (*stems/ha*), age 50	678	37	1,231

densities for *B* and *H* seems to have reversed between ages 12 and 50. Another noteworthy outcome is that while the specified prior mean for *B* of 31.6 m^2/ha was near the center of the age 12 posterior density, the specified prior values for *H* (12.6 *m*) and *N* (1942.4 *stems/ha*) do not appear to be plausible values according to the age 12 posterior densities. On the other hand, the observed means from location 3, which entered through the likelihood (31.5 m^2/ha, 10.9 *m*, 2,101.3 *stems/ha*), all appear to be very reasonable values according to figure 1. This is usual in Bayesian analysis; given a relatively vague prior, the posterior is ordinarily strongly influenced by the data.

The posterior density for number of stems at age 50 is surprising. Of the 5,000 observations in the age 50 sample, 11 percent had fewer than 100 *stems/ha*. This indicates that about one-tenth of the time we can expect very poor survival. Commonly used yield tables predict somewhat greater survival. For instance, the survival model Bailey *et al.* (1985) predicts survival of 335 *stems/ha* at age 50. However, growth and yield models usually are based on data from much younger stands (loblolly pine stands tend to be harvested well before age 50), and published studies show remarkable variation in predicted numbers of surviving stems per unit area (Burkhart *et al.* 1981). Long-term studies with many replications would be required to determine whether our results are reasonable.

More results, including analysis of the influence of the likelihood on ϕ, and examination of the posterior distribution of θ, are included in Green *et al.* (1998).

CONCLUSION

BSYN is a promising technique for generating posterior distributions and/or interval estimates for the outputs from mechanistic models. Unlike other methods of quantifying uncertainty in mechanistic models, BSYN incorporates all the information available to the user about the inputs and outputs of the model, in addition to the model itself. The end result is a joint posterior distribution of the model inputs and outputs, which may be used to make valid inferences about any of the input or output parameters, or any mathematical function of them.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grant #DEB-9602201, a grant from the USDA Forest Service Northern Global Change Program, and New Jersey Agricultural Experiment Station funds. We thank R.L. Amateis and H.E. Burkhart for data contributions to this study, R.L. Bailey and M.A. Buford for help in assembling growth and yield model information, and A.E. Raftery and G.H. Givens for helpful discussions on Bayesian Synthesis.

LITERATURE CITED

Amateis, R.L.; Burkhart, H.E.; Knoebel, B.R.; Sprinz, P.T. 1984. Yields and size class distributions for unthinned loblolly pine plantations on cutover site-prepared lands. Publ. FWS-2-84. Virginia Polytechnic Institute and State University. 69 p.

Bailey, R.L.; Grider, G.E.; Rheney, J.W.; Pienaar, L.V. 1985. Stand structure and yields for site-prepared loblolly pine plantations in the Piedmont and Upper Coastal Plain of Alabama, Georgia, and South Carolina. Res. Bull. 328. University of Georgia, College Agricultural Experiment Station. 118 p.

Burkhart, H.E.; Cao, Q.V.; Ware, K.D. 1981. A comparison of growth and yield prediction models for loblolly pine. Publ. FWS-2-81.Virginia Polytechnic Institute and State University, School of Forestry and Wildlife Resources. 23 p.

Clutter, J.L.; Brister, W.L.; Rheney, J.H. 1984. Stand structure and yields of site-prepared loblolly pine

plantations in the lower Coastal Plain of the Carolinas, Georgia, and North Florida. Gen. Tech. Rep. SE-27. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 173 p.

Givens, G.H.; Bravington, M.V. 1995. Borel's paradox and sensitivity analysis. Pap. SC/47/AS15 presented to IWC Scientific Committee. May 1995.

Givens, G.H.; Roback, P.J. 1997. Logarithmic pooling of priors linked by a deterministic simulation model. Tech. Rep. 97-21. Colorado State University, Department of Statistics. 19 p.

Green, E.J.; MacFarlane, D.W.; Valentine, H.T. 1998. Assessing uncertainty in a stand growth model by Bayesian synthesis. Manuscript submitted to Forest Science.

Hafley, W.L.; Smith, W.D.; Buford, M.A. 1982. A new yield prediction model for unthinned loblolly pine in plantations. Tech. Rep. 1. North Carolina State University, School of Forest Research, Bioeconomical Modeling Project. 65 p.

Johnson, N.L.; Kotz, S. 1970. Continuous univariate distributions - 2. New York, NY: John Wiley and Sons. 306 p.

MacFarlane, D.W.; Green, E.J.; Valentine, H.T. 1998. Parameter distribution estimation for a stand-level carbon balance model for loblolly pine. Manuscript in preparation.

Poole, D.; Raftery, A.E.. 1998. Assessment of the Bering-Chukchi-Beaufort Seas stock of Bowhead whales using Bayesian and full pooling Bayesian synthesis methods. Paper presented to the IWC Scientific Committee. April 1998.

Raftery, A.E.; Givens, G.H.; Zeh, J.E. 1995. Inference from a deterministic population dynamics model for Bowhead whales. JASA 90: 402-442, with discussion and rejoinder.

Raftery, A.E.; Poole, D.; Givens, G.H. 1996. The Bayesian synthesis assessment method: resolving the Borel paradox and comparing the backwards and forwards variants. Pap. SC/48/AS16 presented to the IWC Scientific Committee. June 1996.

Raftery, A.E.; Poole, D. 1997. Bayesian synthesis assessment methodology for Bowhead whales. Pap. SC/49/AS5 presented to the IWC Scientific Committee. September 1997.

Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. In: Bernardo, J.M.; *et al.*, eds. Bayesian Statistics 3. Oxford, UK: Clarendon Press: 395-402.

Schweder, T.; Hjort, N.L. 1996. Bayesian synthesis or likelihood synthesis? What does the Borel paradox say? Report of the International Whale. Commission. 46: 475-480.

Shinozaki, K.; Yoda, K.; Hozumi, K.; Kira, T. 1964a. A quantitative analysis of plant form: the pipe model theory. I. Basic analyses. Japan Journal of Ecology. 14: 97-105.

Shinozaki, K.; Yoda, K.; Hozumi, K.; Kira, T. 1964b. A quantitative analysis of plant form: the pipe model theory. II. Further evidence of the theory and its application in forest ecology. Japan Journal of Ecology. 14: 133-139.

Thornley, J.H.M. 1970. Respiration, growth and maintenance in plants. Nature. 227: 304-305.

Thornley, J.H.M. 1971. Energy, respiration and growth in plants. Annals of Botany. 25: 721-728.

Thornley, J.H.M. 1976. Mathematical models in plant physiology. London: Academic Press.

Valentine, H.T. 1988. A carbon-balance model of stand growth: a derivation employing pipe-model theory and the self-thinning rule. Annals of Botany. 62: 389-396.

Valentine, H.T. 1990. A carbon-balance model of tree growth with a pipe-model framework. In: Dixon, R.K.; *et al.*, eds. Process modeling of forest growth responses to environmental stress. Portland, OR: Timber Press: 33-40.

Valentine, H.T.; Gregoire, T.G.; Burkhart, H.E.; Hollinger, D.Y. 1997. Pipestem: a stand-level model of carbon allocation and growth, calibrated for loblolly pine. Canadian Journal of Forest Research. 27: 817-830.

Wolpert, R.L. 1995. Comment on "Interference from a deterministic population dynamics model for Bowhead whales, by Raftery, A.E.; Givens, G.H.; Zeh, J.E. 1995.JASA 90: 402-442, with discussion and rejoinder.