Finding the appropriate level of complexity for a simulation model: An example with a forest growth model

Rasmus Astrup [a,*], K. David Coates [b], Erin Hall [b]

[a] Bulkley Valley Centre for Natural Resources Research and Management, 1188 Main Street, Box 4274, Smithers, BC V0J 2N0, Canada
[b] British Columbia Forest Service, Research Section, Bag 6000, Smithers, BC V0J 2N0, Canada

ABSTRACT

The topic of model complexity is fundamental to model developers and model users. In this study, we investigate how over- and under-fitting of a driving function in a simulation model influences the predictive ability of the model. Secondly, we investigate whether model selection approaches succeed in selecting driving functions with the best predictive ability. We address these issues through an example with the forest simulator SORTIE-ND. Utilizing maximum likelihood methods and individual tree growth data we parameterize five growth functions of increasing complexity. We then incorporate each growth function into the simulation model SORTIE-ND and test predicted growth against independent data. Compared to the independent data, the simplest and the most complex growth functions had the poorest predictive ability while functions of intermediate complexity had the best predictive ability. The poor predictive ability of the simplest model is caused by poor approximation of the system while the poor predictive ability of the most complex model is caused by biased parameter estimates. A growth function of intermediate complexity was the most parsimonious model where error due to approximation and error due to estimation were simultaneously minimized. The model selection criteria AIC and BIC were found to select complex functions that were over-fitted according to the independent data comparison. BIC was closer to choosing the model that minimized prediction error than AIC. In this example, BIC is the more appropriate model selection criterion. It is important that both model developers and models users remember that more complex models do not always result in better predictive models.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Simulation models are increasingly popular in both ecology and forest management (Messier et al., 2003). There is broad agreement among model developers that parsimonious models are preferable over more complex models (Gauch, 2003; McLeod, 1993; Pidd, 1996). Parsimony has been an important topic for both philosophers of science and applied scientists for several hundred years (Forster and Sober, 1994; Gauch, 2003). The meaning of parsimony is often explained through various versions of Occham's razor: "entities should not be multiplied beyond necessity" (e.g. Hoffmann et al., 1996; Young et al., 1996; Forster, 2000; Steel and Penny, 2000; Gauch, 2003), a supposed Einstein quote: "as simple as possible but no simpler" (e.g. Burnham and Anderson, 2002), or in a more statistical sense as "the smallest possible number of

parameters for adequate representation of data" (Box and Jenkins, 1970).

The objective of most simulation models is to make predictions about a system of interest. The main benefit of utilizing a parsimonious model is that it provides the best predictions (Burnham and Anderson, 2002; Gauch, 1993, 2003). Parsimonious models provide better predictions because the error due to approximation (simplification of the system) and error associated with parameter estimation are balanced to obtain the overall lowest level of prediction error (Box and Jenkins, 1970; Burnham and Anderson, 2002; Gauch, 2003).

Even though the benefits of parsimony are widely recognized, the complexity of the available simulation models is increasing rapidly (Pidd, 1996; Harte, 2002). For most systems and research questions, our knowledge of the system has not increased at the same rate as the complexity of our modeling tools. According to Chwif and Paul (2000) there are several reasons for the observed increase in model complexity: an "include all syndrome" that tempts model developers to include everything they know about a system even though a model should be an abstraction of the

* Corresponding author. Tel.: +1 250 847 2817; fax: +1 250 847 2848.
E-mail address: rasmusastrup@gmail.com (R. Astrup).

system, a "possibility factor" that relates to the accessibility of computing power, and finally "unclear simulation objectives" which cause models to be unnecessarily complex. Alternatively, it is possible that model complexity has increased because the simpler models did not provide adequate answers to our questions. The trend towards higher degrees of model complexity combined with the importance of parsimony raises two main questions: (1) how to achieve the appropriate level of complexity in modeling and (2) what happens to model predictions when models are too complex or too simple.

The way that many ecologists analyze data in order to make an inference to a population has shifted in the last two decades (e.g. Johnson and Omland, 2004). The shift has been from null-hypothesis testing towards a model selection framework (e.g. Burnham and Anderson, 2002; Dayton, 2003; Johnson and Omland, 2004) where multiple working hypotheses (Chamberlin, 1965) represented as mathematical models are compared against data. The best approximating model or set of best approximating models is used to make inference to the population and selected with a model selection criterion. The most frequently used model selection criteria are AIC (Akaike, 1973) and BIC (Schwarz, 1978) (Kuha, 2004).

In ecology and econometrics, AIC or the small sample size equivalent AIC_C (Hurvich and Tsai, 1989) appears dominant while BIC is prominent in social sciences (Kuha, 2004). Heuristic interpretation of most model selection criteria include the concept of parsimony as models are awarded for better fit but penalized for increasing complexity measured by the number of parameters (e.g. Bozdogan, 2000; Burnham and Anderson, 2002; Dayton, 2003). For simulation models that are based on data, alternate formulations of the driving functions can be derived within a model selection framework. In this case, the goal of the model selection is to ensure the selection of functions with the best predictive ability.

In this study, we purposely parameterize five tree growth functions (models) of varying complexity from field data. We then compare each model using two model selection criteria (AIC_C and BIC). We use the forest dynamics simulator SORTIE-ND to evaluate the predictive ability of each model against independent data. Our analysis addresses two important issues of interest to model developers and model users. First, it addresses how complexity (over- or under-fitting) of a driving function in a simulation model influences the predictive ability of the model. The absolute values from this analysis will be specific to the evaluated model and data, but the patterns will be general and valid across a wide range of models and modeling approaches. Second, our analysis addresses whether the use of standard model selection approaches ensures selection of the function with the best predictive ability. The results related to model selection criteria are limited to simulation models where the driving functions are fitted directly from data.

2. Methodology

2.1. Mature tree growth functions

The growth functions used in our analysis predict annual radial increment of individual trees and are variations of the models developed by Canham et al. (2004, 2006). A thorough explanation of the model structure and biological interpretation of the individual parameters can be found in the original papers. In this analysis, we use individual tree growth data collected in stem-mapped plots from forests of the Sub-Boreal Spruce, moist cold subzone (SBSmc2; Banner et al., 1993) located near the town of Smithers in central British Columbia, Canada. We use data for the four dominant tree species in these forests: interior spruce [a complex of white spruce Picea glauca (Moench) Voss and Engelmann spruce (Picea engel-

mannii Parry ex Engelm.)], lodgepole pine (Pinus contorta var. latifolia Engelm.), subalpine fir (Abies lasiocarpa (Hook.) Nutt.), and trembling aspen (Populus tremuloides Michx.).

The utilized growth functions include tree species, tree size, estimated shading, and a distance-dependent competition index (a proxy for belowground competition) as predictor variables (Canham et al., 2004). We investigated a total of five nested models with decreasing levels of complexity and an associated decreasing number of parameters. In the remainder of this paper the five models are referred to as the: Full Model, Crowding–Shading Model, Crowding Model, Shading Model and Size Model. The most complex model is the Full Model and the other four models are simplifications of this model. In the Full Model, the predicted growth of an individual tree is a maximum growth rate (Max Growth) modified by three individual effects: the tree size (Size Effect), shading from neighbours (Shading Effect) and competition from neighbours (Crowding Effect):

$$Growth = Max\ Growth \times Size\ Effect \times Shading\ Effect$$
$$\times Crowding\ Effect,$$

where Growth is the radial increment (mm/year), Max Growth is an estimated parameter that determines the maximum radial increment that can be predicted for an individual tree, and the individual effects have the functional forms outlined in the following equations:

$$Size\ Effect = e^{-0.5[\ln(DBH/X_0)/X_b]^2}, \tag{1}$$

where DBH is the diameter of the target tree in centimeters (at 1.30 m height above ground level) and X_0 and X_b are estimated parameters that describe the mode and variance. As tree size increases, the size effect allows for an initial increase in predicted radial increment, followed by a peak where the Max Growth is obtained (DBH = X_0) and a slow decrease in radial increment as tree diameter increases above DBH = X_0.

$$Shading\ Effect = e^{-m \times S}, \tag{2}$$

where S is the amount of shading from neighbour trees (0 = no shade and 1 = full shade) while m is an estimated parameter that determines the rate of decrease in growth associated with shading.

$$Crowding\ Effect = e^{-C \times DBH^D \times NCI} \tag{3}$$

where DBH is the diameter of the target tree and NCI is a competition index (Neighbourhood Crowding Index (4)) while C and D are estimated parameters. The C parameter allows for a decrease in growth with increased competition and the D parameter allows different tree sizes to have differential sensitivity to competition (Canham et al., 2004). NCI is calculated according to (4) where the crowding is summed across i = 1 to s species and j = 1 to n neighbours, $dist_{ij}$ is the distance to a given $neighbour_{ij}$; α_i and β_i are estimated species-specific parameters; and λ_i is a species-specific competition intensity factor, which ranges from 0 to 1, for each neighbour species relative to the target tree (Canham et al., 2004). Thus, when four species are analyzed, four species-specific estimates of λ_i are obtained for each target tree species.

$$NCI = \sum_{i=1}^{s} \sum_{j=1}^{n} \lambda_i \frac{(DBH_{ij}/100)^{\alpha_i}}{dist_{ij}^{\beta_i}} \tag{4}$$

For each tree species, the Full Model has a total of 13 estimated parameters and includes the Size Effect, the Crowding Effect where different tree sizes are differentially influenced by competition, and the Shading Effect. The Crowding–Shading Model contains 12 parameters and is the second most complicated model. Compared to the Full Model, the Crowding–Shading Model is simplified by

Table 1
Summary of the main effects included in the five nested models

	Size Effect	Shading Effect	Crowding Effect	Differential Crowding Effect	Number of parameters
Full Model	Yes	Yes	Yes	Yes	13
Crowding–Shading Model	Yes	Yes	Yes	No	12
Crowding Model	Yes	No	Yes	No	11
Shading Model	Yes	Yes	No	No	4
Tree Size Model	Yes	No	No	No	3

fixing parameter D (Eq. (3)) in the crowding effect at zero. Thus, in the Crowding–Shading Model, different tree sizes are equally sensitive to competition. The Crowding Model has a total of 11 estimated parameters and is simplified from the Full Model by setting both parameters D in the *Crowding Effect* and parameter m in the shading effect equal to zero. Thus, the Crowding Model represents a model where shading has no effect on predicted radial growth and all tree sizes are equally sensitive to competition. The Shading Model has a total of four estimated parameters and is derived from the Full Model by excluding the *Crowding Effect*. Thus the Shading Model only allows shading and tree size to influence tree growth. The final and most simple of the five tested models is the Size Model. The Size Model has three estimated parameters and is derived from the Full Model by excluding both the *Shading Effect* and the *Crowding Effect*. Thus in the Size Model, growth of each species is exclusively dependent on tree size and independent of the surrounding neighbourhood. A summary of the five utilized models can be found in Table 1.

2.2. Fitting the growth functions

The five growth functions were fitted to individual tree data from 14 stem-mapped stands (averaging 0.3 ha in size) that included a large range of tree sizes (DBH range: 3–65 cm), ages, and species mixtures. The positions of all trees taller than 130 cm were recorded and the DBH of all trees were measured. To get estimates of radial increment, an increment core was taken from a subsample ($n_{Total} = 600$) of lodgepole pine ($n_{Pine} = 135$), subalpine fir ($n_{Fir} = 158$), trembling aspen ($n_{Aspen} = 57$) and interior spruce ($n_{Spruce} = 250$). The five growth functions were fitted to the data by maximum likelihood estimation (e.g. Edwards, 1992) with a normal probability density function (e.g. Hilborn and Mangel, 1997) and utilizing the global optimization method simulated annealing (Goffe et al., 1994). Average radial increment over the past 5 years was the measure of tree growth.

2.3. Estimation of prediction error with independent data

To evaluate the performance of the five individual growth functions we incorporated each function into the individual tree, spatially explicit stand dynamics simulator SORTIE-ND and then compared the SORTIE-ND growth predictions (basal area m^2/ha) to measured growth in 49 permanent sample plots over a 30-year period. The permanent sample plot growth data was obtained through the British Columbia Ministry of Forests long-term database. The permanent plots were either 0.1 or 0.08 ha in size and were established in either 1970 or 1971. All 49 plots were conifer-dominated with either lodgepole pine or interior spruce as the dominant species. Tree species other than interior spruce, lodgepole pine, subalpine fir, and trembling aspen made up a maximum of 4% of the crown cover in the initial year of measurement. The 49 permanent sample plots were from similar site types and age classes as the 14 stem-mapped plots used to parameterize the five growth functions. The initial stand density in the permanent sample plots ranged between 273 and 2645 stems/

ha with a mean density of 1389 stems/ha. The basal area ranged between 10.3 and 54.5 m^2/ha with a mean basal area of 29.7 m^2/ha.

To simulate the permanent sample plots in SORTIE-ND, the tree list from the initial plot measurement (1970/1971) was repeated to create a 9 ha stem-map file that was then input into SORTIE-ND. Individual trees in each stem-map file were assigned random (x,y) locations by SORTIE-ND. For fully stocked stands in the sub-boreal forests of British Columbia, unpublished model tests have shown that random locations for the individual trees are an appropriate starting condition for SORTIE-ND simulations.

The growth of all 49 permanent sample plots was simulated with each of the five growth functions. The predicted stand basal areas were then compared to the measured growth of the permanent sample plots. The accuracy of the predictions for each plot was described by the mean actual residual (bias) calculated with formula (5). The precision was described by calculating the mean absolute residual calculated with formula (6). These two performance measures are commonly used in the evaluation of forest growth models (Vanclay and Skovsgaard, 1997).

$$\text{mean actual residual} = \sum_{r=1}^{r=k} \frac{\hat{y}_r - y_r}{k} \quad (5)$$

$$\text{mean absolute residual} = \sum_{r=1}^{r=k} \frac{|\hat{y}_r - y_r|}{k} \quad (6)$$

where $\hat{y}_r$ is the predicted increment for plot r, y_r is the observed increment for plot r and k is the total number of plots.

3. Results

In general, both AIC$_C$ and BIC decreased with increasing model complexity for the three conifer species resulting in one of the two most complex models always having the lowest AIC$_C$ and BIC values (Table 2 and Fig. 1). For trembling aspen growth, AIC$_C$ and BIC both indicated similar levels of support for the simpler Shading and Size Models (Table 2). Neither model selection criteria suggested a more complex model for aspen. BIC appeared to favour slightly simpler models than AIC$_C$ (Table 2).

For interior spruce, BIC suggested similar levels of support for the three most complex models with the Crowding Model having the least support of the three. AIC$_C$ indicated similar levels of support for the interior spruce Full and Crowding–Shading Models and no support for other models (Table 2). Both AIC$_C$ and BIC clearly selected the Full Model for lodgepole pine growth. Results for subalpine fir were similar to those for lodgepole pine except the BIC method also showed support for the Crowding Model (Table 2).

Similar R^2-values were observed for the Full Model, the Crowding–Shading Model, and the Crowding Model by tree species. The R^2-values for the three models were within a range of 0.04 within a species, while the range among species was 0.56–0.87 (Table 2 and Fig. 1). The R^2-values for the Shading Model and the Size Model were much lower than for the other models (Table 2 and Fig. 1).

Table 2
AIC_C and BIC for the five mature tree growth functions

Model	Interior spruce	Lodgepole pine	Subalpine fir	Trembling aspen
Full	**ΔAIC_C = 0.0 (1)** ΔBIC = 2.2 (2) R^2 = 0.66	**ΔAIC_C = 0.0 (1)** **ΔBIC = 0.0 (1)** R^2 = 0.87	**ΔAIC_C = 0.0 (1)** **ΔBIC = 0.0 (1)** R^2 = 0.74	ΔAIC_C = 18.5 (5) ΔBIC = 29.1 (5) R^2 = 0.56
Crowding–Shading	ΔAIC_C = 1.1 (2) **ΔBIC = 0.0 (1)** R^2 = 0.67	ΔAIC_C = 7.3 (2) ΔBIC = 4.9 (2) R^2 = 0.86	ΔAIC_C = 18.4 (3) ΔBIC = 6.8 (3) R^2 = 0.70	ΔAIC_C = 11.2 (4) ΔBIC = 21. 2 (4) R^2 = 0.59
Crowding	ΔAIC_C = 7.3 (3) ΔBIC = 3.22 (3) R^2 = 0.65	ΔAIC_C = 25.6 (3) ΔBIC = 30.8 (3) R^2 = 0.84	ΔAIC_C = 16.3 (2) ΔBIC = 2.0 (2) R^2 = 0.70	ΔAIC_C = 8.4 (3) ΔBIC = 17.9 (3) R^2 = 0.59
Shading	ΔAIC_C = 149.2 (4) ΔBIC = 118.0 (4) R^2 = 0.34	ΔAIC_C = 136.7 (4) ΔBIC = 111.2 (4) R^2 = 0.59	ΔAIC_C = 78.3 (4) ΔBIC = 41.6 (4) R^2 = 0.50	**ΔAIC_C = 0.0 (1)** ΔBIC = 0.5 (2) R^2 = 0.46
Size	ΔAIC_C = 240.9 (5) ΔBIC = 206.2 (5) R^2 = 0.04	ΔAIC_C = 221.9 (5) ΔBIC = 193.6 R^2 = 0.21	ΔAIC_C = 183.7 (5) ΔBIC = 144.1 (5) R^2 = 0.01	ΔAIC_C = 1.2 (2) **ΔBIC = 0.0 (1)** R^2 = 0.43

Values in bold represent the models selected by AIC_C and BIC. The numbers in parentheses are the ranking from 1 to 5.

In no cases would AIC_C have selected the Crowding Model for any of the four tree species. It was 7.3–25.6 units greater than the lowest AIC_C value for the Full Model (Table 2). The BIC method indicated limited support for the Crowding Model for subalpine fir only (2 units lower than the best model).

When the individual growth models were incorporated into the SORTIE-ND simulation model and tested against independent data, the Crowding Model provided the best approximation to the permanent sample plot data while the Full Model resulted in the poorest approximation (Table 3 and Fig. 1). The most complex model with the most parameters did not provide the best predictions for the permanent sample plot data (Fig. 1). All five models had positive mean actual residual and positive mean absolute residual and their values were almost identical (Table 3). This indicates that all five models consistently over-predict the growth of the permanent sample plots. The high degree of similarity between the two types of residuals suggests that basal area was under-predicted in only a few of the permanent plots.
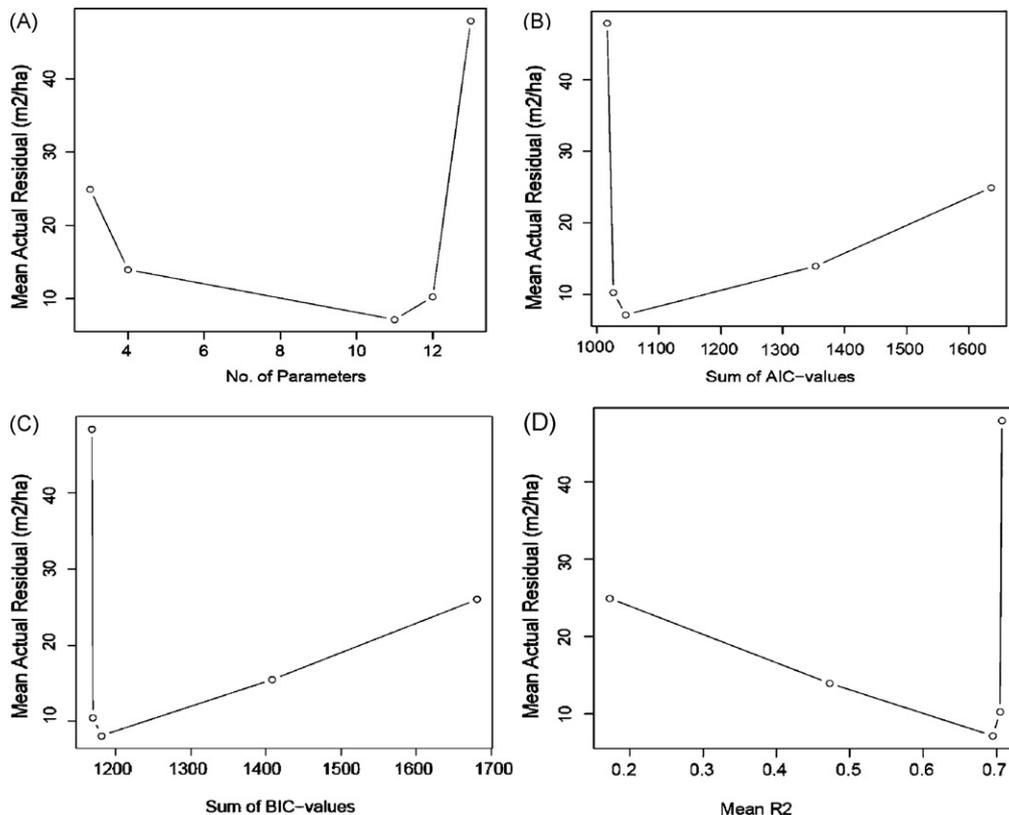

Fig. 1. Mean absolute residual from 49 permanent sample plots as a function of the properties of the mature tree growth function. Mean absolute residual plotted against: (A) the number of parameters/species in the growth function; (B) AIC_C summed across the four species; (C) BIC summed across the four species; and (D) mean R^2 for the four species-specific models.

Table 3
Summary statistics for the 30-year simulation of 49 permanent sample plots

Response variable	Year	Mean observed value	Presence in no. of plots	Mean actual residual	Mean absolute residual
Full Model	10	34.2	49	16.4	18.7
	20	37.7	49	32.8	34.0
	30	38.8	49	47.9	48.4
Crowding–Shading Model	10	34.2	49	3.56	3.78
	20	37.7	49	6.38	6.74
	30	38.8	49	10.2	10.4
Crowding Model	10	34.2	49	2.10	4.06
	20	37.7	49	4.89	5.91
	30	38.8	49	7.17	8.11
Shading Model	10	34.2	49	3.93	4.97
	20	37.7	49	8.14	9.72
	30	38.8	49	13.9	15.5
Size Model	10	34.2	49	6.94	7.71
	20	37.7	49	15.0	16.0
	30	38.8	49	24.9	26.0

A negative mean actual residual indicates that basal areas predicted by the model were on average lower than the plot values. The mean actual residual value illustrates the average difference between the predicted basal areas and the observed basal areas.

4. Discussion

We utilized the spatially explicit forest growth model SORTIE-ND to investigate the relationship between model complexity and predictive ability. There are two levels to determining the appropriate complexity for a simulation model. The higher level is related to the overall objectives of the model. Here, the main concern is that the model structure matches the modeling objectives (e.g. Ward, 1989; Pidd, 1996; Burkhart, 2003). The second level of model complexity, which is the main topic of this paper, is whether the complexity of the individual functions applied within the overall model type and structure is appropriate to the available data.

Our analysis illustrates problems with both under- and over-fitting of relationships in simulation models. This can be best demonstrated by looking at the Size Model and the Full Model. The Size Model was clearly under-fitted and represents an over-simplification of the system. This model omits key processes that influence individual tree growth (see Canham et al., 2004; Stadt et al., 2007) resulting in a model with poor predictive ability (Fig. 1). The main flaw of the Size Model is that growth of an individual tree was considered independent of the surrounding trees which results in prediction errors when densities or sizes of neighbouring trees change.

The Full Model includes more key processes than the Size Model, but in the comparison to the permanent sample plots, the Full Model performs worse than all other models, including the oversimplified Size Model. The main flaw of the Full Model is related to parameter D in (3) which allows different trees sizes to have differential sensitivity to competition. The species-specific parameter estimates for the D parameter were approximately −0.5 for all species which indicates that large trees are quite insensitive to competition. The poor predictive ability of the Full Model (Fig. 1) illustrates that the D parameter estimates are biased and cannot be generalized from the sample to the population at large (the permanent sample plots). If the parameter estimates for D had been approximately zero (equivalent to the Crowding–Shading Model) the predictive ability would have been better (Fig. 1).

The Full Model was over-fitted and it was not the lack of key processes (as in the Size Model) but problems with estimation of the parameter values that resulted in the poor predictive ability of the Full Model.
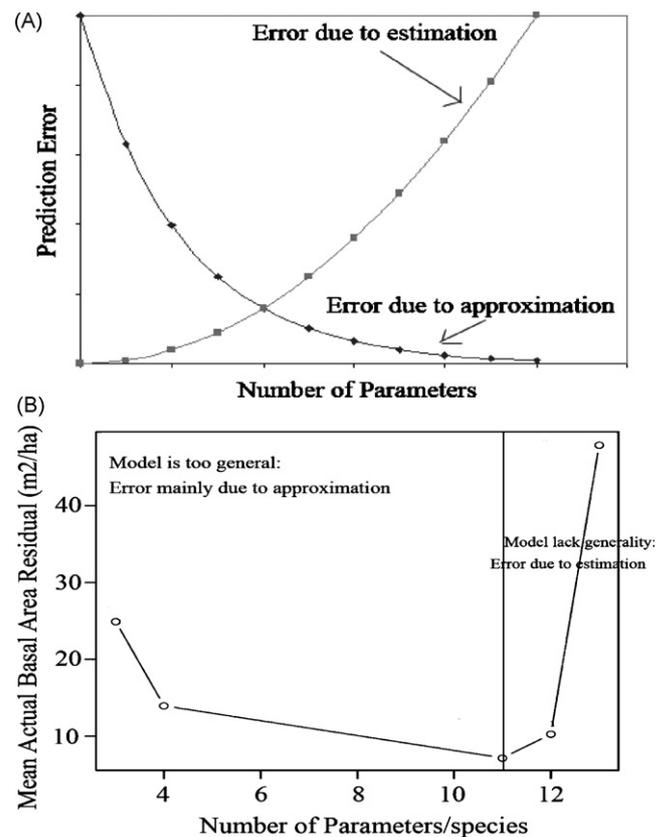

Fig. 2. Explaining the prediction error. (A) A theoretical illustration of error at different levels of model complexity (modified from Burnham and Anderson, 2002) and (B) explaining the shape of the curve in Fig. 1 as error due to approximation and error due to estimation.

Through the process of model selection and model testing we found an appropriate level of model complexity for predicting tree growth in SORTIE-ND given our data availability. If the parameterization dataset had been larger, a more complex model would likely have been the best predictive model. In the same way, other forest simulation models have a complexity level that optimizes the predictive ability dependent on data availability. The more general and important result, which is valid across simulation model approaches and types, is the pattern in the results. The observed pattern of high prediction error for simple models caused by poor approximation of the system (oversimplification) and poor parameter estimation for complex models (Fig. 2A) is well known (e.g. Box and Jenkins, 1970; Burnham and Anderson, 2002; Gauch, 2003), however, actual examples where researchers have independent data and have taken the opportunity to illustrate this pattern for simulation models are rare. The best predictive model should simultaneously minimize error caused by approximation of the system and error caused by estimation of parameters. Selecting a parsimonious model is crucial to achieving good predictive ability. This was clearly demonstrated in our comparison of growth rate predictions by the five different growth functions in the SORTIE-ND model simulations to the measured growth rates in actual forests (permanent sample plot data). In our comparative study, the most parsimonious model was the Crowding Model (Fig. 2B). The more complex models were too dataset specific and had poor predictive ability mainly due to estimation error, while the more simple models were too general and the error was mainly caused by oversimplification of the system.

Most researchers do not have the benefit of a simulation model and independent data to test different functions that represent alternate hypotheses of system behavior. Thus, the problem is choosing which function to utilize (Forster, 2000). A common, easily applicable, and holistic approach is to use a model selection criterion (e.g. Burnham and Anderson, 2002; Dayton, 2003; Johnson and Omland, 2004). We calculated the widely utilized model selection criteria AIC_C and BIC and found that, in general, both criteria favoured more complex models than the Crowding Model that we found to be most parsimonious by evaluation against independent data. With one exception, the two model selection techniques favoured the Full Model, yet this model had poor predictive ability. The Full Model was favoured by the model selection techniques because the data showed a strong tree-size dependent sensitivity to competition which was accommodated by the Full Model. However, our independent dataset indicate that this strong tree-size dependent sensitivity was dataset specific and did not apply to the population at large. In a larger dataset, we assume that the parameter for the tree-size dependent sensitivity to competition would be estimated at a more realistic value.

Our results should not be seen as a flaw in the model selection approach or criteria. It is impossible to create a technique that safeguards against the unfortunate combination of a strong but non-general trend in data and a model created to accommodate this effect or a generally poor model formulation. The risk of either problem occurring can be minimized by avoiding data dredging and data mining and by careful a priori model formulation (Burnham and Anderson, 2002). However, it still occurred in our analysis even though we carefully avoided these known pitfalls.

Our results suggested that BIC was closer to choosing the model that minimized prediction error than AIC_C (Table 3). The literature contains several studies that compare the performance of AIC, AIC_C and BIC and depending on the application either criterion has been shown to be "best" (Huang and Dayton, 1995; Lin and Dayton, 1997; Burnham and Anderson, 2002, 2004; Gagne and Dayton, 2002; Dayton, 2003; Kuha, 2004).

AIC/AIC_C and BIC have different assumptions (Kuha, 2004) and are appropriate in different situations (Burnham and Anderson, 2004). BIC approaches its target model from below and for common (smaller) sample sizes generally favours less complex models than AIC/AIC_C (Burnham and Anderson, 2004). AIC/AIC_C should generally be used in scenarios with many tapering effects that all contribute strongly to the overall system behavior while BIC is appropriate in the case of a few major effects (often represented as nested models) and small tapering effects (Burnham and Anderson, 2004). There are many factors that influence tree growth, but for healthy trees on a given site there are really only a few major factors that influence growth. This can be seen in the change in R^2-values illustrated in Fig. 1D. Our analysis is an example of a scenario where BIC should be applied rather than AIC/AIC_C.

In summary, our results illustrate the importance of complexity in modeling. To achieve good predictive ability, it is important to use a parsimonious model. Model selection techniques can aid in the selection of the most parsimonious model but consideration should be given to using the most appropriate model selection criterion. Additionally, model selection does not prevent a poor model or the combination of a strong data-specific effect and a model that accommodates this to be selected. To assess predictive ability we maintain that as often as possible models should be tested against independent data. There is a growing tendency for simulation models to become more complex. It is important for both model developers and models users to remember that more complex models do not always result in better predictive models. We generally know more about a system than we have data to include in a model.

References

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second international symposium on information theory, Akademiai kiado.

Banner, A., MacKenzie, W., Haeussler, S., Thomson, S., Pojar, J., Trowbridge, R., 1993. A field guide to site identification and interpretation for the Prince Rupert Forest Region. B.C. Ministry of Forests, Research Branch, Victoria, B.C. Land Management Handbook No. 26.

Box, G., Jenkins, G., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Bozdogan, H., 2000. Akaike's information criterion and recent developments in information complexity. Journal of Mathematical Psychology 44, 62–91.

Burkhart, H.E., 2003. Suggestions for choosing an appropriate level for modelling forest stands. In: Amaro, A., Reed, D., Soares, P. (Eds.), Modelling Forest Systems. CAB International, Wallingford, pp. 3–10.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference, a Practical Information—Theoretic Approach, 2nd Edition. Springer-Verlag, New York, Inc..

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods & Research 33, 261–304.

Canham, C.D., LePage, P., Coates, K.D., 2004. A neighborhood analysis of canopy tree competition: effects of shading versus crowding. Canadian Journal of Forest Research 34, 778–787.

Canham, C.D., Papaik, M.J., Uriarte, M., McWilliams, W.H., Jenkins, J.C., Twery, M.J., 2006. Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. Ecological Applications 16, 540–554.

Chamberlin, T.C., 1965. The method of multiple working hypotheses. Science 148, 754–759 (reprint of 1890 paper in Science).

Chwif, L., Paul, R.J., 2000. On simulation model complexity. In: Joines, J.A., Barton, R.R., Kang, K., Fishwick, P.A. (Eds.), 2000 Winter Simulation Conference. pp. 449–455.

Dayton, C.M., 2003. Model comparison using information measures. Journal of Modern Applied Statistical Methods 2, 281–292.

Edwards, A.W.F., 1992. Likelihood. John Hopkins University Press, Baltimore, MD.

Forster, M., 2000. Key concepts in model selection: performance and generalizability. Journal of Mathematical Psychology 44, 205–231.

Forster, M., Sober, E., 1994. How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. British Journal of Philosophical Science 45, 1–35.

Gagne, P., Dayton, C.M., 2002. Best regression model using information criteria. Journal of Modern Applied Statistical Methods 1, 479–488.

Gauch, H.G., 1993. Prediction, parsimony and noise. American Scientist 81, 468–478.

Gauch, H.G., 2003. Scientific Method in Practice. Cambridge University Press.

Goffe, W.L., Ferrier, G.D., Rogers, J., 1994. Global optimization of statistical functions with simulated annealing. Journal of Econometrics 60, 65–99.

Harte, J., 2002. Toward a synthesis of the Newtonian and Darwinian worldviews. Physics Today 55, 29–34.

Hilborn, R., Mangel, M., 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton, New Jersey.

Hoffmann, R., Minkin, V.I., Carpenter, B.K., 1996. Ockham's razor and chemistry. Bulletin de la Societe Chimique de France 133, 117–130.

Huang, C.-C., Dayton, C.M., 1995. Detecting patterns of bivariate mean vectors using model-selection criteria. British Journal of Mathematical & Statistical Psychology 48, 129–147.

Hurvich, C.M., Tsai, C.-L., 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307.

Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19, 101–108.

Kuha, J., 2004. AIC and BIC: comparisons of assumptions and performance. Sociological Methods & Research 33, 188–229.

Lin, T.S., Dayton, C.M., 1997. Model-selection information criteria for non-nested latent class models. Educational and Behavioral Statistics 22, 249–264.

McLeod, A.I., 1993. Parsimony, model adequacy and periodic correlation in time series forecasting. International Statistical Review 61, 387–393.

Messier, C., Fortin, M.-J., Schmiegelow, F., Doyon, F., Cumming, S., Kimmins, J.P., Seely, B., Welham, C.J.N., 2003. Modelling tools to assess the sustainability of forest management scenarios. In: Burton, P.J., Messier, C., Smith, D.W., Adamowicz, W.L. (Eds.), Towards Sustainable Management of the Boreal Forest. National Research Council of Canada. NRC Research Press, Ottawa, pp. 531–580.

Pidd, M., 1996. Five principles of modelling. In: Charnes, J.M., Morrice, D.J., Brunner, D.T., Swain, J.J. (Eds.), 1996 Winter Simulation Conference. pp. 721–728.

Schwarz, G., 1978. Estimating the dimension of a model. Annals of Statistics 6, 461–464.

Stadt, K.J., Huston, C., Coates, K.D., Feng, Z., Dale, M.R.T., Lieffers, V.J., 2007. Evaluation of competition and light estimation indices for predicting diameter growth in mature boreal mixed forests. Annals of Forest Science 64, 477–490.

Steel, M., Penny, D., 2000. Parsimony, likelihood, and the role of models in molecular phylogenetics. Molecular Biology and Evolution 17, 839–850.

Vanclay, J.K., Skovsgaard, J.P., 1997. Evaluating forest growth models. Ecological Modelling 98, 1–12.

Ward, S.C., 1989. Arguments for constructing simple models. Journal of Operational Research Society 40, 141–153.

Young, P., Parkinson, S., Lees, M., 1996. Simplicity out of complexity in environmental modelling: Occam's razor revisited. Journal of Applied Statistics 23, 165–210.