Ecology, 80(2), 1999, pp. 538–553

MULTI-CRITERIA ASSESSMENT OF ECOLOGICAL PROCESS MODELS

JOEL H. REYNOLDS[1,3] AND E. DAVID FORD[2]

[1]*Quantitative Ecology and Resource Management Program, University of Washington, Seattle, Washington 98195 USA*
[2]*Quantitative Ecology and Resource Management Program, and College of Forest Resources, Box 352100,*
University of Washington, Seattle, Washington 98195 USA

Abstract. The Pareto Optimal Model Assessment Cycle (POMAC), a multiple-criteria model assessment methodology, is described for exploring uncertainty in the relationships between ecological theory, model structure, and assessment data. Model performance is optimized to satisfy, simultaneously, each component of a vector of assessment criteria (model outputs), rather than the usual procedure of optimizing performance with respect to a single criterion. Pareto Optimality is used to define the vector optimization. The Pareto Optimal Set reveals which combinations of assessment criteria the model can satisfy simultaneously. Binary interval error measures, which classify whether a parameterization result is within an acceptable range of values, are defined for each criterion. Their use masks small differences in the performance of different parameterizations, allowing the Pareto Optimal Set to reveal conflicts in ability to achieve simultaneously different collections of criteria.

POMAC improves the researcher's ability to detect deficiencies and locate their sources. It is more stringent and informative than traditional model assessment procedures because it uses multiple criteria without weighting and aggregating them. The Pareto Optimal Set reveals the presence of deficiencies through the model's inability to satisfy all the criteria simultaneously. POMAC then guides the researcher in locating deficiencies in: inadequate selection of component ecological hypotheses underlying the model, inadequate mathematical representations of these hypotheses, inadequate parameterization, poor selection and formulation of the assessment criteria, or combinations of these. In an example, POMAC is applied to the spatially explicit canopy competition model WHORL using ten assessment criteria. Each criterion was selected to provide information on different aspects of WHORL's functioning: three stand height distribution criteria, three crown morphology criteria, and four criteria focusing on stand competition's characteristic differentiation of growth rates. The Pareto Optimal Set was generated using simulated evolution optimization. POMAC revealed deficiencies in both the model structure and its assessment criteria, leading to an improved model and better understanding of its effective domain.

Key words: binary error measures; canopy competition; evolutionary computation; genetic algorithms; individual based models; model assessment; model validation; multiple criteria; Pareto Optimality; POMAC; simulated evolution; WHORL.

INTRODUCTION

Complex simulation models of ecological processes are increasingly constructed for use in both the development of ecological theory (e.g., Pacala and Deutschman 1995, Fitz et al. 1996) and the analysis of environmental questions (e.g., Ågren et al. 1991, Landsberg et al. 1991, McMurtrie and Comins 1996, Schimel et al. 1997). Such models can never be validated due to the limited observation of system dynamics (Oreskes et al. 1994, Rykiel 1996). They can, however, be assessed to investigate deficiencies in the relationships they define between ecological theory, model structure, and assessment data. We propose a methodology for this assessment.

There are four potential deficiency sources in an ecological process model, each associated with a different phase of the modeling activity: inadequate selection of the component ecological hypotheses (an incorrect process structure), inadequate mathematical representation of these hypotheses (an incorrect mathematical structure), inadequate fitting procedure (a faulty parameterization), and inadequate selection and formulation of the assessment criteria (an insufficient model assessment context). An assessment methodology must be capable of both detecting each type of deficiency and of guiding the researcher to the source of each deficiency.

Assessments using only a single criterion have limited ability to detect deficiencies because the criterion may be satisfied by many different model structures. This is the problem of nonuniqueness: models with different hypotheses or different mathematical representations of the same hypotheses can satisfy the same criterion equally well. Nonuniqueness prevents the pos-

---

Manuscript received 17 December 1996; revised 25 November 1997; accepted 19 January 1998.

[3] Present address: Department of Statistics, Box 35-4322, University of Washington, Seattle, Washington 98195 USA.

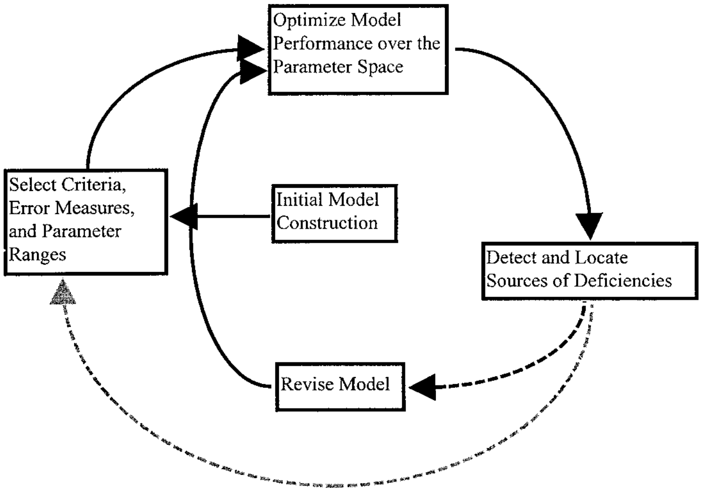

FIG. 1. General model assessment cycle. Having completed initial model construction (center box), the assessment cycle begins with the selection and formulation of the criteria, their error measures, and the parameter space search. Model performance is optimized with respect to the selected criteria in order to reveal deficiencies in either the model structure (black dashed arrow) or the criteria formulations (gray dashed arrow). Either type of deficiency requires revision (dashed lines). If no deficiencies are detected, the criteria and error measures may be refined for a more stringent assessment. The assessment cycle is then repeated to investigate the revisions.

sibility of ultimate model validation (Oreskes et al. 1994). Single criterion assessments, however, are currently the most common type of assessment applied to ecological process models. For example, the canopy competition models in a recent review were each assessed using a single criterion (Ford and Sorrensen 1992).

Using multiple criteria to assess a model increases the demands on model structure. The model is required to get more right. Simultaneous multiple-criteria assessments have a higher capacity to detect model deficiencies than do any single-criterion, or even sequential multiple-criteria, assessments. This benefit has been discussed in the context of individual-based models (Gross et al. 1992, Murdoch et al. 1992, DeAngelis and Rose 1992, Sorrensen-Cothern et al. 1993), but no methodologies have been proposed to utilize these models' many outputs, i.e., potential criteria (Gross et al. 1992).

In ecological modeling, the most common method of multiple criterion assessment is to calculate error measures of different model outputs (criteria) and collapse those measures into a single criterion (e.g., by weighting and summing) (e.g., Halfon 1979, Gentil and Blake 1981, Beck 1987, Sievanen et al. 1988, Sorrensen-Cothern et al. 1993). This single criterion is then optimized by a parameter space search. While such an aggregated measure of model performance simplifies the optimization task, it is exactly the model's simultaneous performance on these different measures (criteria) that needs to be observed in order to locate deficiency sources. Aggregation into a single criterion can also limit the detection of deficiencies by limiting direct observation of the model's performance.

We introduce a multiple-criteria assessment methodology, the Pareto Optimal Model Assessment Cycle (POMAC), which retains the multiple criteria as a vector rather than aggregating them into a single criterion. Model performance is assessed by first optimizing the criteria vector, using the Pareto Optimality definition of vector optimization (see *Pareto Optimization* section). This optimization reveals the combinations of criteria that the model can simultaneously satisfy. The inability of the model to simultaneously satisfy all the criteria directly reveals deficiencies. The Pareto Optimal Model Assessment Cycle then proceeds through four stages (see *The Pareto Optimal Model Assessment Cycle* section, below) to determine the type of each deficiency and locate its source. Having located a deficiency, its source is revised and then the assessment repeated to check for improvement and further model deficiencies (Beck 1985, 1987, Beck and Halfon 1991) (Fig. 1).

We describe Pareto Optimization and discuss its use in deriving a multiple-criteria summary of model performance. The Pareto Optimal Model Assessment Cycle is then described in six stages and demonstrated in application to the assessment of the canopy competition model WHORL (Sorrensen-Cothern et al. 1993). The potential use of POMAC in comparing different models and assessing aggregation of a large model is discussed.

PARETO OPTIMIZATION

The Pareto Optimality definition of vector optimization (Vincent and Grantham 1981) was first developed in economics and is used mainly there and in engineering (e.g., Taylor et al. 1975, Olenik and Haimes 1979, Vincent 1987). To select the optimum from a collection of vectors of criteria results, each from a different model parameterization, one compares the vectors and removes all dominated ones, where vector **X** dominates vector **Y** if and only if **X** is at least as good as **Y** with respect to all criteria and there is at least one criterion for which **X** is strictly better than **Y**.

The Pareto Optimal Set consists of those vectors left after all dominated ones have been removed. Most frequently, the Pareto Optimal Set is used as a basis for constructing additional multiple-criteria optimization techniques through various methods of weighting and aggregating the criteria into a single criterion (see Steuer 1986 or Yu 1985). Here we use it as a basis from which to build a model assessment methodology.

For example, assume three criteria have been selected to judge different aspects of a canopy competition model for trees: cumulative mortality in the stand, median live tree height, and mean individual tree crown depth as proportion of tree height. Each criterion is calculated at the end of a fixed simulation time span corresponding to a time interval for which data are available. Each parameterization produces a vector having three components, one component for each of the three criteria. For each criterion a binary interval is defined covering an acceptable range of expected results, so that a prediction falling within the interval is considered acceptable with regard to this criterion and is labeled ''good''; otherwise it is labeled ''bad.'' For example, cumulative mortality predictions within the range [77, 137] are ''good,'' predictions outside this range are ''bad.'' Suppose the simulations of four different parameterizations result in the following assessment vectors, with components mortality, median height, and crown depth:

Assessment vector for parameterization 1
= (bad, good, bad)
Assessment vector for parameterization 2
= (bad, good, good)
Assessment vector for parameterization 3
= (good, bad, good)
Assessment vector for parameterization 4
= (good, bad, good).

Parameterization 2 is as good as parameterization 1 with regard to mortality and median height while actually doing better with regard to crown depth; thus, parameterization 2 dominates parameterization 1. Parameterizations 3 and 4 neither dominate, nor are dominated by, parameterizations 2 or 1. No judgment can be made regarding precedence among parameterization 2 and parameterizations (3, 4) without resorting to weighting or preference among the criteria (e.g., mortality is more important than median height, so parameterizations 3 and 4 both dominate parameterization 2). The Pareto Optimal Set is the set of undominated parameterizations, i.e., parameterizations {2, 3, 4}.

The binary interval error measures do not discriminate among parameterizations that generate the same assessment vector. For example, while parameterizations 3 and 4 produce different criterion values, these values generate identical assessment vectors. The Pareto Optimal Set is summarized by these parameterization groups rather than by the individual parameterizations themselves. The Pareto Optimal Set above has two groups: parameterization 2 and parameterizations (3, 4) (Table 1).

TABLE 1. Example of a Pareto Optimal Set derived from the four hypothetical tree-canopy competition model parameterizations discussed in *Pareto Optimization.*

Group	Parameterizations	Assessment vector		
		Mortality	Median ht.	Crown depth
1	2		X	X
2	3, 4	X		X

Notes: Three criteria are used: cumulative stand mortality, median live tree height, and mean individual tree crown depth as proportion of tree height. Each row presents a group of parameterizations, which produced a common assessment vector (last three columns). The parameterizations in each group are listed in the ''Parameterizations'' column. An "X" in a cell denotes that each parameterization in the group adequately simulates the criterion; an empty cell denotes that no parameterization in the group satisfied the criterion.

In practice, the Pareto Optimal Set is achieved when sufficient parameterizations of the model have been simulated to explore the parameter space and ensure that no additional combinations of assessment vectors can occur. Note that the Pareto Optimal Set need not necessarily contain a parameterization achieving all criteria: not only may the model require different parameterizations to satisfy different criteria (e.g., Table 1), it may never satisfy specific individual criteria or combinations of criteria.

Binary interval error measures

Binary interval error measures ignore small differences between the simulations of different parameterizations, treating results close to a criterion's proposed target value uniformly (Hornberger and Spear 1981, Hornberger and Cosby 1985, Jaffe et al. 1987). This robustness to possible target value misspecification is essential when the criterion is designed to capture an ecological process whose specific target value is not certain. In contrast, traditional error functions (e.g., least squares, etc.) are designed to be sensitive to a criterion's proposed target value, and misspecification can heavily influence the optimization search. Similarly, these traditional error measures, when used in a multiple-criteria optimization that aggregates the criteria to a single measure, may restrict the optimization search too quickly to parameter regions that produce near-optimal results for a specific criterion (Kursawe 1991). This restriction may prevent the search from locating parameterizations achieving better overall satisfaction of the collected criteria—exactly the parameterizations of interest in model assessment. Binary error measures are robust to both of these concerns.

Binary error measures have been employed in sensitivity analysis (Hornberger and Cosby 1985, Mäkelä 1988), to investigate uncertainty in model forecasts (Rose et al. 1991, van Straten and Keesman 1991), and to predict the order of time series models (Keesman

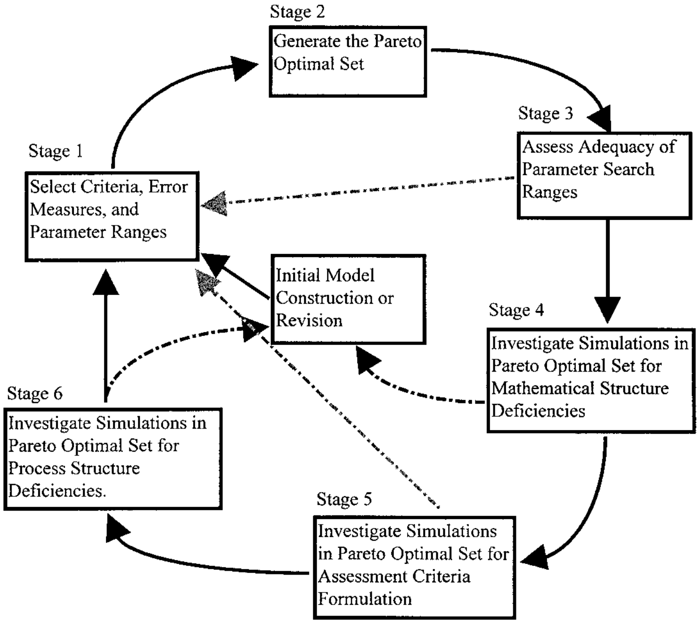

FIG. 2. The Pareto Optimal Model Assessment Cycle, a refinement of the process illustrated in Fig. 1. Having completed initial model construction (center box), the assessment context is defined (Stage 1). The model's Pareto Optimal Set is generated by an optimization search over the selected parameter space (Stage 2). Deficiencies located in Stages 3–6 require revision and reinitiation of the cycle (dashed arrows); otherwise assessment proceeds to the next stage (solid arrows). Deficiencies located in Stages 3 or 5 require revision of the assessment context (gray dashed arrows); deficiencies located in Stages 4 or 6 require revision of the model structure (black dashed arrows). If no deficiencies are located by the end of Stage 6, the binary error intervals can be made more restrictive, or more refined criteria can be selected and the cycle reinitiated.

and van Straten 1989). To the best of our knowledge, they have not been employed with a vector of criteria, nor used to aid the summary of a Pareto Optimal Set.

THE PARETO OPTIMAL MODEL ASSESSMENT CYCLE

The Pareto Optimal Model Assessment Cycle (POMAC) expands the general model assessment cycle (Fig. 1) to include a stage of investigation specifically focusing on each of the four potential sources of model performance deficiency (Fig. 2): incorrect process structure, incorrect mathematical structure, faulty parameterization, or insufficient model assessment context.

The initial stages of POMAC define the assessment criteria and the parameter space search (Fig. 2, Stage 1), and generate the Pareto Optimal Set (Fig. 2, Stage 2). Assessment then proceeds by first checking to ensure that deficiencies revealed in the Pareto Optimal Set are not arising from too limited a parameter search range (Fig. 2, Stage 3). Next the simulation results of representative parameterizations from each group are investigated for deficiencies in the model's mathematical structure (Fig. 2, Stage 4) and in assessment criteria formulations (Fig. 2, Stage 5). Any remaining deficiencies can be attributed to the model's process structure (Fig. 2, Stage 6). Each stage is demonstrated in application to the spatially explicit canopy competition model WHORL (Sorrensen-Cothern et al. 1993).

Initial model construction: the tree canopy competition model WHORL

WHORL is a spatially explicit model of competition among trees. Two stages are reiterated annually: (i) resource (light) distribution within the stand and acquisition by individuals, depending upon their amount and distribution of foliage, and (ii) growth and resource allocation within individuals, producing new foliage (Sorrensen-Cothern et al. 1993). It simulates competition in *Abies amabilis,* a shade tolerant species, in clustered natural regeneration at a density of 90 000 trees/ha growing to a height of 7 m. WHORL has two simulation modes: growth of a stand of competing trees and growth of a single tree in a neighborless environment (open-grown). Trees are composed of modules, i.e., branches that grow and die independently. The volume within which the stand grows consists of ''cells'' (cubes 10 cm on a side). A branch's resource acquisition occurs in those cells it intersects in which it supports foliage. Different branches may extend into the same cell. A cell's foliage density changes, branches grow, and trees increase in height according to the amount of resource acquired. Physiological plasticity in resource acquisition and allocation is represented by allowing differences in foliage characteristics, e.g., radiation interception efficiency (Table 2), as a function of a tree's relative height.

Stage 1: Selecting the assessment context

Parameter search ranges.—The parameter search ranges for WHORL were based on prior calibrations of the model (Sorrensen-Cothern et al. 1993, Reynolds 1997) (Table 2).

Criteria selection.—Assessment criteria differ in their sensitivities to the model parameters. Ten criteria were chosen so that both modes of model simulation, open-grown and crown competition, were assessed for each parameterization. Six criteria measure general stand and open grown crown characteristics (Table 3,

TABLE 2. Parameter search ranges for the spatially explicit tree-canopy competition model WHORL.

Parameters	Description	Minimum	Maximum	Scale
Branch increment	Branch growth rate (m/unit of relative production)	0.0001	0.0007	0.0001
Height increment	Tree height increment rate (m/unit of relative production)	0.00002	0.00010	0.00001
Dead	Minimum production required to sustain a branch (units of relative production)	0.0	1.0	0.05
E	Efficiency of converting intercepted irradiation to production (relative scale of production units per unit of irradiation flux)	1.0	3.0	0.5
D	Minimum relative irradiation level required to sustain living foliage (percentage of full irradiation level)	0	12	1
K	Exponent controlling nonlinear decline in foliage response to irradiation decay (dimensionless)	0.1	0.5	0.1

Notes: See Sorensen-Cothern et al. (1993) for a more detailed description of the parameters. Scale refers to the parameter-specific minimum step size used in the optimization search. Relative production is production per unit foliage, an index of production in excess of maintenance respiration of leaves (Sorrensen-Cothern et al. 1993). Physiological plasticity is represented by varying the values of a tree's physiological parameters (Dead, *E*, *D*, and *K*) as a function of the tree's height class. Height classes are defined relative to the fifth tallest live tree: short = height $\leq 0.55 \times$ reference tree height; medium $= 0.55 \times$ reference height $<$ tree height $\leq 0.75 \times$ reference height; tall = height $> 0.75 \times$ reference height. This requires a Short, Medium, and Tall parameter setting for Dead, *E*, *D*, and *K*, for each simulation.

stand and open-grown criteria), while the remaining four focus on the specific differentiation in growth rates commonly observed in a stand of trees competing for light (Table 3, growth rate criteria). It was expected that the more specific growth rate criteria would be harder for a model to satisfy than the more general stand and open-grown tree criteria.

Error measures.—Binary interval error measures were selected for each criterion based on data from a permanent plot (Table 3) (see Sorrensen-Cothern et al. 1993 for plot details). The goal was to investigate if the model could produce accurate simulations rather than precisely recreate the exact values observed. Error intervals were constructed from either 95 or 99% confidence intervals, centered on observed data, where possible. Researchers familiar with the species provided binary error intervals for the three open grown criteria (T. Hinckley and R. Brookes, *personal communication*).

Parameter space search technique.—An evolutionary simulation optimization routine searched WHORL's parameter space and generated its Pareto Optimal Set (see Stage 2). Simulated evolution is a suite of optimization techniques that simulate natural selection in order to evolve optimal parameterizations (Michalewicz 1992, Fogel 1994). These include Genetic Algorithms (Goldberg 1989, Holland 1992), Evolutionary Strategies (Bäck et al. 1991), Evolutionary Programming (Fogel 1994), and their extensions (Michalewicz et al. 1992). The terms used are borrowed from genetics. Each simulation from a ''population'' of parameterizations is assessed and ''parent parameterizations'' are then selected using a ''fitness'' measure. The search's next ''generation'' of parameterizations are ''bred'' from the parents' population either through parameter ''mutation,'' i.e., random selection and adjustment of a specific setting, or ''crossover recombination,'' i.e., the exchange of portions of the settings of two parent parameterizations.

This technique is not required for generating the Pareto Optimal Set, but is more efficient than a simple lattice search (Reynolds 1997). The Pareto Optimal Set cannot be generated using ''hill-climbing'' techniques (Yu 1985), either deterministic or stochastic (e.g., Simulated Annealing [Uhry 1989]), as these require a continuous univariate cost function. Evolutionary optimization programs require only that a procedure be defined for selecting parameter values for the next generation's parents using a measure based on criterion achievement. Multiple criteria can be successfully used in a vector form, even with binary error functions, by allowing membership in the Pareto Optimal Set to be used as the optimization's measure of achievement (Reynolds 1997).

Stage 2: Generating the Pareto Optimal Set

The evolutionary optimization routine used both single parameter mutations (within a given search range) and crossover recombination of two parameterizations. A parameterization was selected for breeding based on a fitness function incorporating both the number of criteria the parameterization satisfied and membership in the Pareto Optimal Set (Reynolds 1997). The optimization routine's initial population consisted of a Pareto Optimal Set generated from a preliminary forward search, augmented with randomly selected parameterizations for genetic diversity (see Reynolds 1997).

Simulated evolution produces successive popula-

TABLE 3. Criteria chosen for WHORL's assessment focus on three different aspects of the simulation: stand height and mortality, morphological characteristics of open-grown tree crowns, and the differentiation in growth rates observed in stands competing for light.

General focus	Criterion	Description	Binary error interval (observed)
Stand criteria	Mortality	Cumulative Mortality, year 30 (number of dead trees)	[77, 137] (107)
	Stand height frequency distribution	*P* value of the Kolmogorov-Smirnov two-sample test comparing predicted and observed live height frequency distributions	[0.01, 1]
	Median live tree height	A robust measure of tree height central tendency	[2.82, 3.35] (3.08 m)
Open-grown criteria	Number of live whorls	The number of living whorls on a 30-yr-old open grown tree	[9, 17] (13)
	Crown angle	Angle formed between the mainstem of the tree and the tips of live branches	[10, 15] degrees
	Crown length ratio	Ratio of live crown length to tree height	[0.90, 1]
Growth rate criteria	Suppressed tree growth rate	Slope estimate from linearly regressing two-year height increment on height, at year 28, of trees $\leq$ 2.8 m (Fig. 4)	[0.014, 0.046] (0.03 m/m)
	Suppressed tree R^2	The variability in suppressed tree height increment rates, measured by the coefficient of determination from the regression above (Fig. 4)	[0.04, 0.44] (0.24)
	Dominant tree slope	Slope estimate from linearly regressing 2-yr height increment on height, at year 28, of trees $\geq$ 3.2 m	[−0.005, 0.085] (0.04 m/m)
	Dominant tree R^2	The variability in suppressed tree height increment rates, measured by the coefficient of determination from the regression above	[0.0, 0.27] (0.07)

Notes: The binary error interval associated with each criterion displays the range of results considered to have adequately simulated the characteristic. For example, the mortality criterion is satisfied by any simulation producing a total mortality of 77–137 trees in the first 30 yr of stand development. Where available, the value of the criterion observed at the permanent plot is given in parentheses below the binary error interval. The interval for median live tree height is a 95% confidence interval for the observed median height; the intervals for the growth rate criteria are approximate 99% confidence intervals constructed from the observed regression estimates ± 3 × standard error of estimate.

tions of model parameterizations. The Pareto Optimal Set was updated after each population's simulation by comparing assessment criteria achievement relative to that of the previous Pareto Optimal parameterizations. The parent population, from which the next offspring parameterizations were bred, consisted of all parameterizations in the current Pareto Optimal Set supplemented with non-Pareto Optimal parameterizations from the last generation's offspring to ensure ''genetic'' diversity. The parent population was of fluctuating size, though always >75 parameterizations. Simulated evolution continued for 50 generations with an offspring population of 75 simulations each generation. At that point the Pareto Optimal Set was considered stable because the assessment vectors had not changed for 13 generations. The Pareto Optimal Set contained 181 different parameterizations, partitioned by their assessment vectors into eight groups. For example, 65 parameterizations produced simulations classified (by the binary error intervals) as achieving the first eight criteria listed in Table 3 but not the last two (Group 1, Table 4).

This computationally intensive stage of the model assessment cycle may limit the size of model to which POMAC can be applied when using simulated evolution.

Stage 3: Assessing the adequacy of the parameter search ranges

WHORL's model structure could satisfy every criterion, but not all ten simultaneously, indicating that deficiencies existed (Table 4). The first potential deficiency source to be checked was the fitting procedure: perhaps the search ranges used in the optimization had been too restricted and should be expanded? Each parameter's search range was examined by plotting the values of the Pareto Optimal parameterizations against the search ranges selected in Stage 1 (Fig. 3). If all parameterizations in the Pareto Optimal Set use a value for a particular parameter at an extreme of its search

TABLE 4. Pareto Optimal Set for the spatially explicit canopy competition model WHORL.

		Assessment vector									
Group	N_P†	Mortality	Ht. distribution	Median ht.	No. live whorls	Crown angle	Crown ratio	Suppressed slope	Suppressed R^2	Dominant slope	Dominant R^2
1	65	X	X	X	X	X	X	X	X		
2	18	X			X	X	X	X	X	X	
3	1	X			X	X	X			X	X
4	39	X	X	X		X	X	X	X	X	
5	1	X			X		X		X	X	X
6	53	X	X	X	X	X		X	X	X	
7	2	X	X		X	X	X			X	
8	2	X			X	X	X		X	X	

Notes: The 181 different parameterizations are partitioned into eight groups, with each parameterization in a group producing the same assessment vector. An X denotes that every parameterization in the group adequately simulates the assessment criterion; no X denotes that no parameterization in the group adequately simulates the criterion. The number of parameterizations in a group depends on the sensitivity of the model to different parameters, the width of the criteria error bounds, and the stochastic nature of the simulated evolution optimization search.

† Number of parameterizations.

range, then the range may have been too limited. If so, the search range must be extended and a new Pareto Optimal Set generated (Fig. 2). This was not the case here (Fig. 3), suggesting that WHORL's deficiencies arose from one of the other deficiency sources: its process structure, its mathematical structure, or the criteria formulations.

Stage 4: Investigating the Pareto Optimal Simulations for mathematical structure deficiencies

As the parameter search ranges appeared adequate, focus moved to WHORL's mathematical structure (Fig. 2) as a possible source of deficiencies (Table 4). The Pareto Optimal Set had captured the most informative model parameterizations, reducing the number of simulations needing investigation from 3750 (the number undertaken in the optimization search) to 181. These were thoroughly investigated, with specific attention to whether the criteria were satisfied in acceptable ways (Fig. 4).

In parameterization Groups 2, 3, 5, 7, 8, and, to a lesser degree, in Groups 4 and 6 (Table 4) simulated tree heights had clustered distributions not found in measured data (Fig. 4). These clusters biased the regression used to determine the dominant tree slope and dominant tree R^2 criteria (for example, Fig. 4, Groups 3 or 7; criteria defined in Table 3). To investigate this effect, regressions were recalculated with clusters removed. Then only parameterizations from Groups 2 and 3 achieved the dominant tree slope criterion, and only one parameterization from Group 3 satisfied the dominant R^2 criterion. The membership in the Pareto Optimal Set of the parameterizations in Groups 4–8 (Table 4) was the result of bias in the regression introduced by the clusters.

The clustering occurred at the tree height class boundaries, which are set each year based on the stand's height distribution (as described in Table 2; see Sorrensen-Cothern et al. 1993). The boundaries group trees into small, medium, and tall height classes. These classes determine the physiological parameter values attributed to a tree for the ensuing growth period (see Table 2 legend); they are unrelated to the height boundary the growth rate criteria use to classify a tree as suppressed or dominant at year 28 (growth rate criteria, Table 3).

Clustering occurred when both parameters Dead and *E* changed markedly between height classes. For example, consider a simulation with Dead parameter values (10, 10, 40) for small, medium, and tall height classes, respectively, and *E* values (2.5, 2.0, 1.0). As the stand grew, a tree slightly taller than the medium/tall height class boundary incurred high branch maintenance costs, i.e., its Dead parameter value was 40 relative production units. Due to competition, the tree fell back into the moderate height class, decreasing its branch maintenance costs by a factor of 4, from 40 to 10 relative production units, and doubling its foliage's conversion efficiency, from 1 to 2. These changes in the tree's parameter values increased its height increment to be larger than the smaller trees of the height class it had just left (Fig. 4 and Group 5). Consequently, it was reclassified into the tall height class the following year—albeit, again, as one of the shortest trees in the tall class. The cycle repeated, ''attracting'' trees to heights around the class boundaries across which Dead or *E* changed markedly. This inadequate representation of physiological plasticity by height class-dependent parameter values had to be revised. After the revision, the Pareto Optimal Assessment Cycle was re-initiated (Fig. 2).

We concluded that representing physiological plasticity on a tree-to-tree scale was too restrictive, producing extreme changes in physiological characteristics year-to-year, even if the tree's local light environment did not change markedly (as the height of the tallest trees determined the height class boundaries). Rather than assigning values for the foliage property

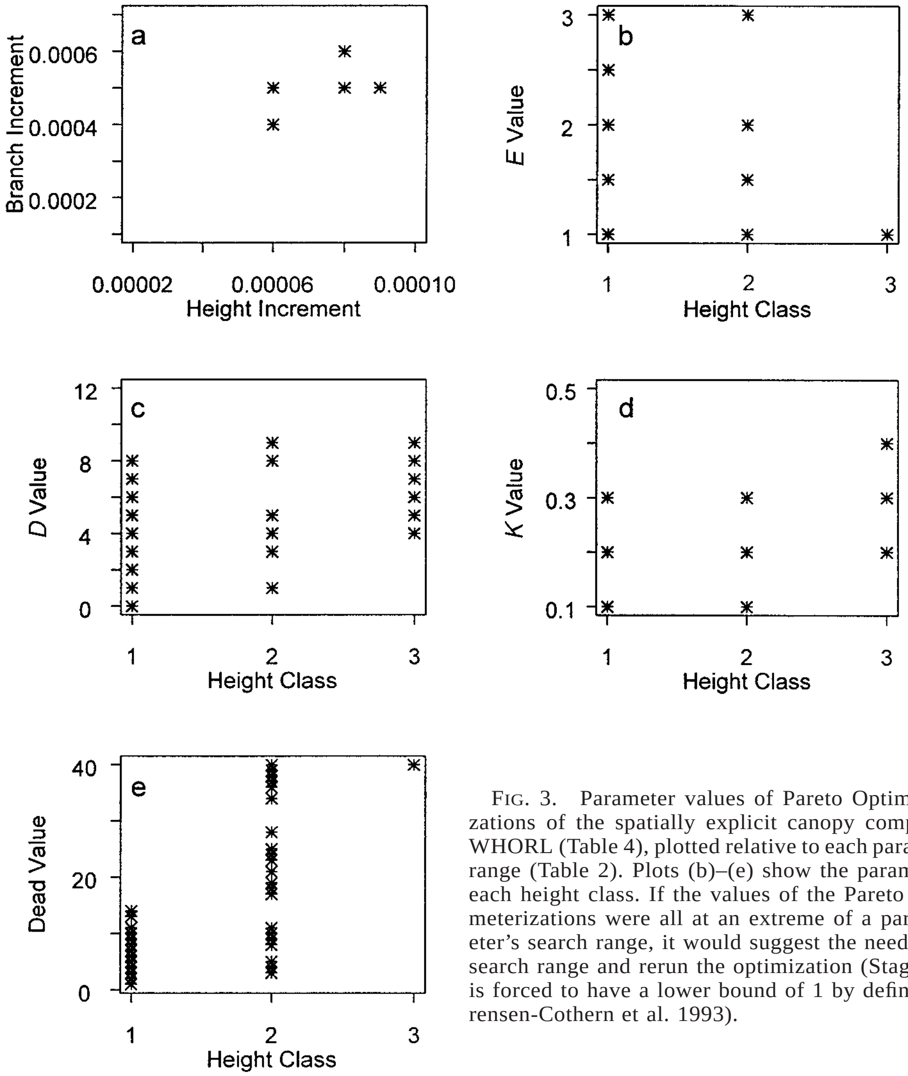

FIG. 3. Parameter values of Pareto Optimal parameterizations of the spatially explicit canopy competition model WHORL (Table 4), plotted relative to each parameter's search range (Table 2). Plots (b)–(e) show the parameter value for each height class. If the values of the Pareto Optimal parameterizations were all at an extreme of a particular parameter's search range, it would suggest the need to expand the search range and rerun the optimization (Stage 2, Fig. 2). *E* is forced to have a lower bound of 1 by definition (see Sorrensen-Cothern et al. 1993).

parameters (*E, D, K*) to a whole tree based on the tree's relative height class in the stand, these parameters would be assigned to each of the tree's foliage cells (a cube 10 cm on a side) according to their illumination level. Each foliage parameter took one of two settings, a shade foliage setting or a sun foliage setting, with a third parameter determining the light intensity at which the setting switched. Foliage properties could differ within and among branches in a tree. The branch maintenance cost parameter, Dead, was made to vary as a simple linear function of branch length with a minimum cost threshold (Fig. 5). The branch and height increment rate parameters remained set at the level of the whole tree and remained fixed for the stand. These changes in physiological plasticity representation were incorporated to form a new model, WHORL2.

The Pareto Optimal Model Assessment Cycle was re-initiated with WHORL2 (Fig. 2). Its Pareto Optimal Set was generated using the simulated evolution optimization routine (Table 5). The Pareto Optimal parameter search ranges were sufficient (Reynolds 1997), and deficiencies did not appear in the model's mathematical structure (Fig. 6).

WHORL2 performed better than WHORL. The parameterizations in Groups 2 and 3 of WHORL2's Pareto Optimal Set (Table 5) simultaneously achieved the general stand criteria (Median Live Height, Mortality) and the more specific dominant growth rate criteria. This combination was not achieved by any parameterization in WHORL's Pareto Optimal Set, represented by Groups 1–3 (Table 4) after removal of the clustering-induced bias.

The continued inability to satisfy all ten criteria simultaneously (Table 5) resulted from deficiencies in either the model's process structure or in the criteria formulations. Achieving the stand height and open-grown criteria (columns 3–10, Table 5) still conflicted with achieving the dominant tree growth rate criteria

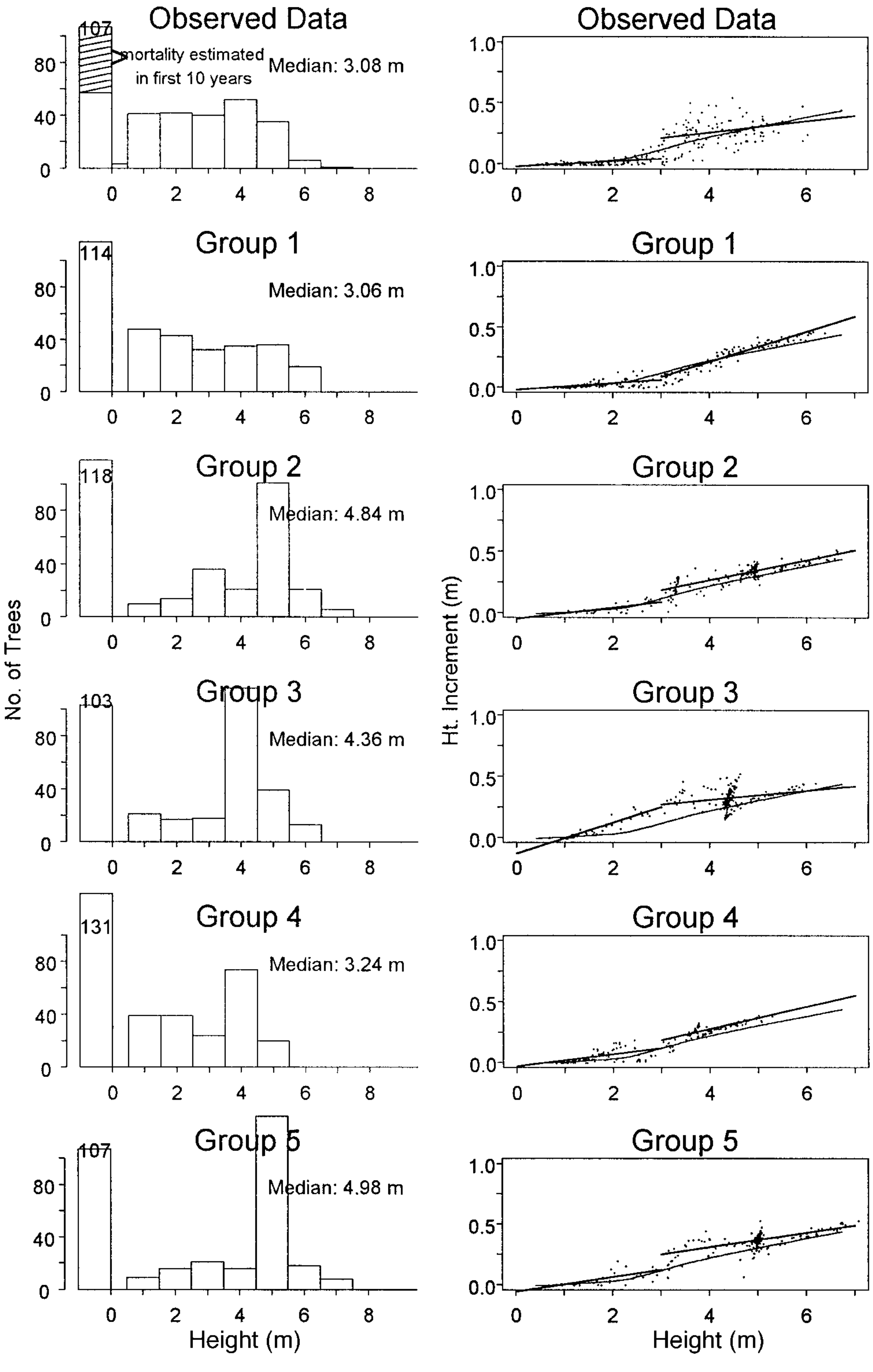

FIG. 4. Simulation results from a representative parameterization from each of the eight groups in WHORL's Pareto Optimal Set (Table 4): tree height frequency distribution at year 30 (left column) and relationship of tree height at year 28 to subsequent 2-yr height increment (right column). The permanent plot (observed) data are shown in the first row of graphs. The leftmost column of each histogram represents cumulative mortality at year 30. Two-year height increment plots display the suppressed and dominant tree growth rate regressions for the simulation (line segments) (Table 3) and, for reference, a locally weighted least-squares smooth (Cleveland 1979) of the observed data (continuous lines). Note the clustering of tree heights in Groups 2–8.

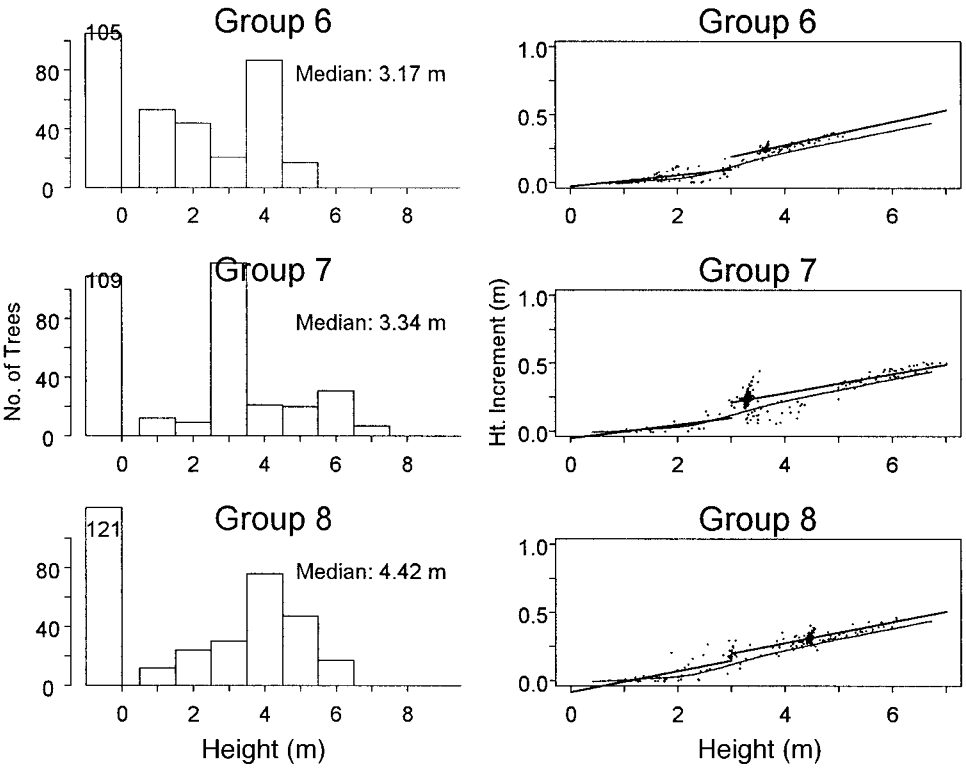

FIG. 4. Continued.

(columns 11, 12, Table 5). Assessment progressed to investigating deficiencies in the criteria formulations and the model's process structure (Fig. 2).

Stage 5: Investigating the Pareto Optimal Simulations for criteria formulation deficiencies

Stage 5 detects simulations achieving criteria in unacceptable ways due to poor criteria selection or formulation. Formulation of a criterion may be too rigid, involving hidden assumptions that produce biased results, or may fail to capture the intended phenomenon.

For example, it was expected that achievement of the four growth rate criteria would require more model refinement than achieving the six stand and open-grown criteria. However, parameterizations in WHORL2's Pareto Optimal Set (Table 5) achieved the more specific dominant tree slope criterion only when failing to achieve all six less specific criteria (Groups 2–8), raising suspicions that the simulations had satisfied the dominant tree slope criterion in an unacceptable manner. This could occur due to deficiencies in the model's mathematical structure, such as the clustering-induced bias in growth rate results found in Stage 3, or in the criteria formulations themselves.

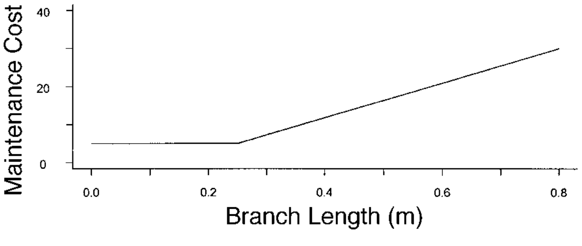

FIG. 5. Revised branch maintenance cost function in the spatially explicit canopy competition model WHORL in relative production units, defined by three parameters: minimum maintenance cost, slope, and intercept.

Suppressed and dominant tree classifications used in the growth rate criteria were defined relative to suppressed and dominant tree height ranges in the observed data (Fig. 6): suppressed trees have heights ≤2.8 m at 28 yr of age, and dominants have heights ≥3.2 m. The impact of using these fixed height ranges to define dominant and suppressed trees was investigated by refitting each Pareto Optimal simulation's growth rate regressions to the simulation's own apparent suppressed and dominant tree height ranges. Under this revision, the dominant tree R^2 criterion was still not satisfied by any of the parameterizations (results ranged from $R^2 = 0.79$ to 0.93 vs. the value of $R^2 = 0.07$ for the observed data) (Reynolds 1997). The dominant tree growth rate slope criterion was satisfied only by the Group 8 parameterization (slope = 0.081 m increment per meter of height), the slopes of all the other parameterizations were too large (0.089–0.149 m/m vs. the observed 0.04 m/m).

The presence in WHORL2's Pareto Optimal Set of the parameterizations in Groups 2–7 was an artifact of

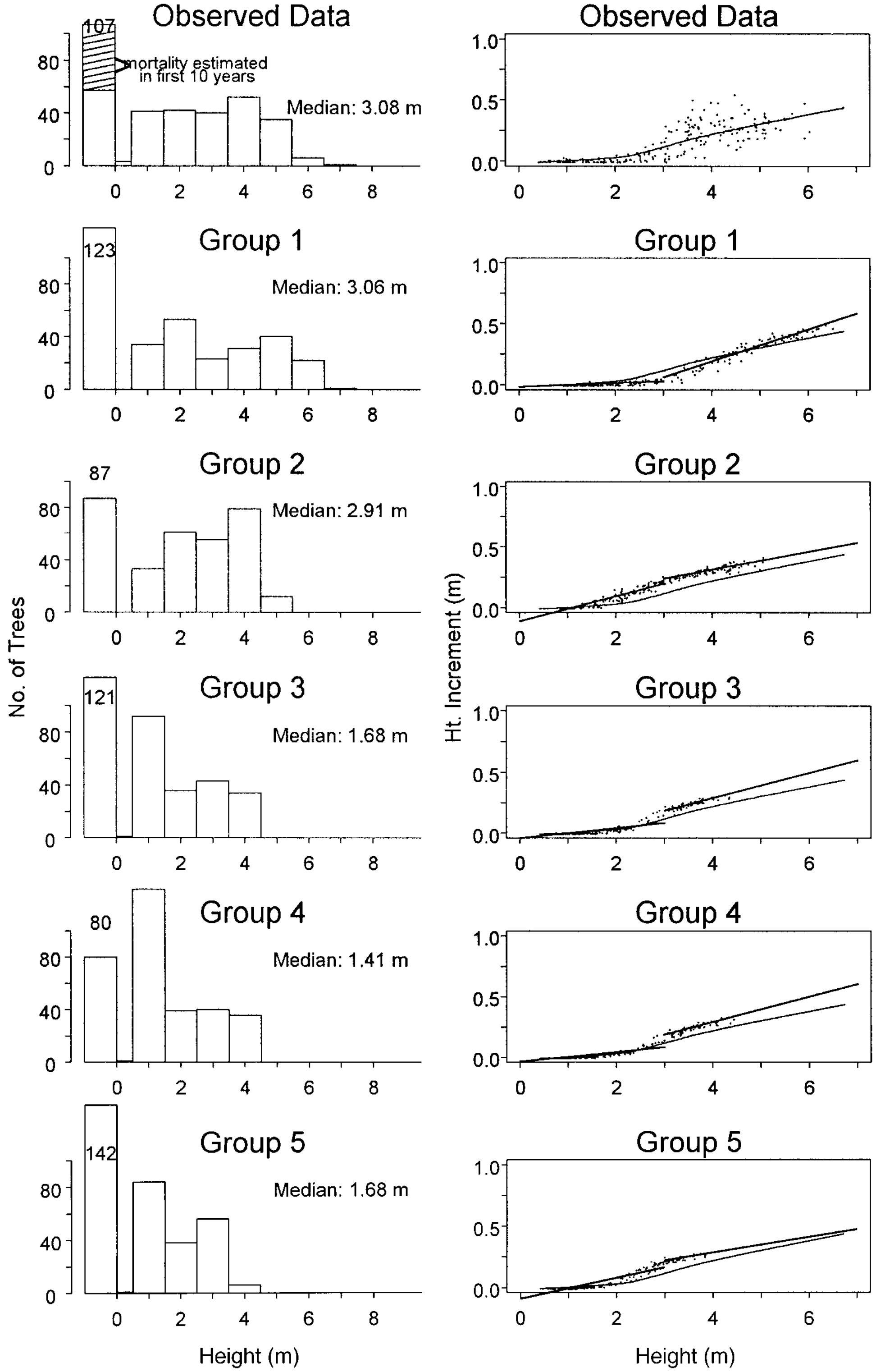

FIG. 6. Simulation results from a representative parameterization from each of the eight groups in WHORL2's Pareto Optimal Set (Table 5): tree height frequency distribution at year 30 (left column) and relationship of tree height at year 28 to subsequent 2-yr height increment (right column). The permanent plot data are shown in the first row of graphs. The leftmost column of each histogram displays cumulative mortality at year 30. The 2-yr height increment plots display the suppressed and dominant tree growth rate regressions for the simulation (line segments) and, for reference, a locally weighted least squares smooth (Cleveland 1979) of the observed data (contiuous lines).

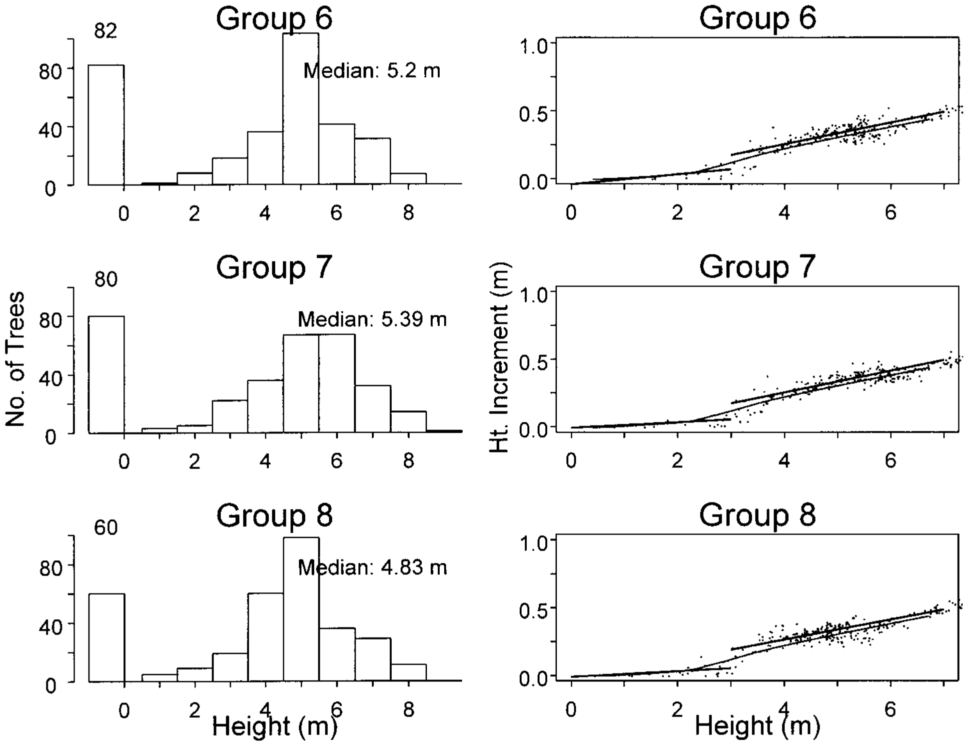

FIG. 6. Continued.

the assumption in the growth rate criteria that the transitional height between suppressed and dominant trees at year 28 would be the same as that in the observed data, i.e., 3 m (Table 3). The dominant growth rate regression only used trees with heights ≥3.2 m. Hence only a subset of the dominant trees in each simulation in Pareto Optimal Groups 2–7 were used in calculating the criteria results, introducing a bias (Fig. 6). Future model assessments should use the simulation's transitional height as an assessment criterion, with the growth rate regressions defined relative to this predicted breakpoint.

The criteria formulations also assumed a linear growth rate response to tree height. While this simplified the criteria calculations, it poorly captured the observed plateau in height increments (Fig. 6, observed height increment plots). At this age, the data indicate that height growth of large trees varies round a common mean rather than continuing to increase with height; the model treats competition as the dominant process influencing growth for all trees.

The assessment process also revealed the inadequacy of the dominant tree R^2 criterion. Variability in dominant tree growth rates is likely due to a variety of factors not included in the model: genotypic variability, microclimate effects, spatial and temporal heterogeneity in soil properties, water, and possibly other factors. In fact, the domain of WHORL2 is the competition process, and any achievement of this dominant tree growth rate criterion should have generated skepticism.

TABLE 5. WHORL2's Pareto Optimal Set (274 different parameterizations). Otherwise, display format as is in Table 4.

		Assessment vector									
Group	N_P†	Mortality	Height distribution	Median height	No. live whorls	Crown angle	Crown ratio	Suppressed slope	Suppressed R^2	Dominant slope	Dominant R^2
1	260	X	X	X	X	X	X	X	X		
2	1	X	X	X		X				X	
3	2	X	X		X	X				X	
4	4	X			X	X		X		X	
5	2	X			X	X				X	X
6	1	X			X				X	X	
7	3	X					X	X	X	X	
8	1				X		X	X	X	X	

† Number of parameterizations.

The dominant tree R^2 criterion was only satisfied as a result of WHORL's clustering or WHORL2's biased regressions, and in retrospect both it and the dominant tree slope criterion were inappropriate selections.

WHORL2 was designed as a competition model, and its ability to simultaneously achieve all the criteria except those with deficient formulations showed it to be effective for this goal (Group 1, Table 5). It effectively simulated the relative reduction in growth rates of small trees, as well as the principal features of stand mortality, stand structure, and individual crown structure. Modeling the absolute growth rate, particularly of dominant trees, would require research and modeling in a different domain.

The choice of binary interval error measures influenced the composition of the Pareto Optimal Set by determining which parameterizations satisfied each criterion. Investigating the impact on the Pareto Optimal Set of small decreases in the binary intervals' widths revealed the sensitivity of the Pareto Optimal Set to particular criteria. The changes from decreasing the width of binary intervals were explored by plotting the Pareto Optimal simulation results against each criterion's error interval (Fig. 7). Slight reductions in the binary intervals of any of five criteria would eliminate two or more groups of parameterizations from the Pareto Optimal Set (Table 6). This highlighted the tenuous nature of membership in WHORL2's Pareto Optimal Set for many of the groups containing small numbers of parameterizations (Table 5). In each case, the eliminated parameterizations' original membership in the Pareto Optimal Set had depended on their achievement of the dominant tree slope criterion (in most cases due solely to the biased criterion formulation discovered above [see *The Pareto Optimal Model Assessment Cycle: Stage 5*). Minor reduction in the dominant tree slope criterion's interval eliminated six of the Pareto Optimal Set's eight parameterization groups (Table 6). This criterion's role in determining WHORL2's Pareto Optimal Set illustrated that, in addition to understanding how a model functions, it is equally important to learn how it can best be assessed.

Investigating slight increases in the binary intervals' widths would reveal whether simultaneous satisfaction of all the criteria was narrowly missed. This would require repeating the evolutionary optimization with the larger intervals. We did not take this next step as WHORL2 had already demonstrated the ability to satisfy all criteria in the domain of focus.

Revising the collection of criteria and their error measures, either by redefining them or by adding or removing criteria, is cause to re-initiate the Pareto Optimal Model Assessment Cycle (Fig. 2). Knowing the inadequacies in the criteria formulations revealed in Stage 5, we continued to Stage 6 to investigate deficiencies in the model's process structure.

Stage 6: Investigating the Pareto Optimal Groups for process structure deficiencies

Deficiencies not due to the parameter search process (Stage 3), the model's mathematical structure (Stage 4), or the criteria selection and formulation (Stage 5), must arise from the model's process structure (Fig. 2). Insight into WHORL2's process structure deficiencies was provided by examining how the parameterizations in each group of the Pareto Optimal Set failed to satisfy their unachieved criteria. Following Stage 5, the dominant tree R^2 criterion was discounted, as was the importance of parameterizations in Groups 2–7 because they only appeared in the Pareto Optimal Set due to bias in the dominant tree growth rate criterion.

Group 8 failed to achieve the general stand and crown angle criteria because of an extremely high height increment parameter setting (Reynolds 1997), i.e., the simulated stand was too tall (Fig. 6, Group 8). Group 1 failed to achieve the dominant tree growth rate criterion, which was actually outside the domain of competition, because its tallest trees grew too fast (Fig. 6, Group 1, height increment plot; Fig. 7, dominant slope). This result suggested that WHORL2's process structure was missing a limit, or control, to keep the growth rate of the tallest trees within the range observed in the permanent plot data. This effect could have resulted from a lack of mainstem maintenance costs, foliage age, or other factors that would progressively diminish the resource capture and utilization capacities of the foliage. If further research were conducted into factors controlling the growth rate of large trees, and additional functions were incorporated into WHORL2, then the assessment cycle should be reiterated, employing new criteria to assess the revised model structure (Fig. 2).

DISCUSSION

Each iteration of the Pareto Optimal Model Assessment Cycle increased insight into the model's capabilities and limitations, increasing the model's value as a heuristic. Deficiencies in the model structure or criteria formulations were deficiencies in our understanding of the phenomenon being modeled. In revealing these deficiencies, POMAC guided and directed further research and data collection.

The application demonstrates that WHORL2 is successful in simulating competition for light as measured by its ability (Table 5) to satisfy simultaneously the first eight criteria in Table 3, and that it is an improvement over WHORL. WHORL2's domain limits are revealed by its inadequate representation of the processes controlling the dominant tree growth rate (Table 5).

Multiple-criteria assessment using Pareto Optimization is a stringent technique for detecting deficiencies. For example, note that none of the deficiencies detected by POMAC would have been revealed by using any single assessment criterion (Table 5), or even

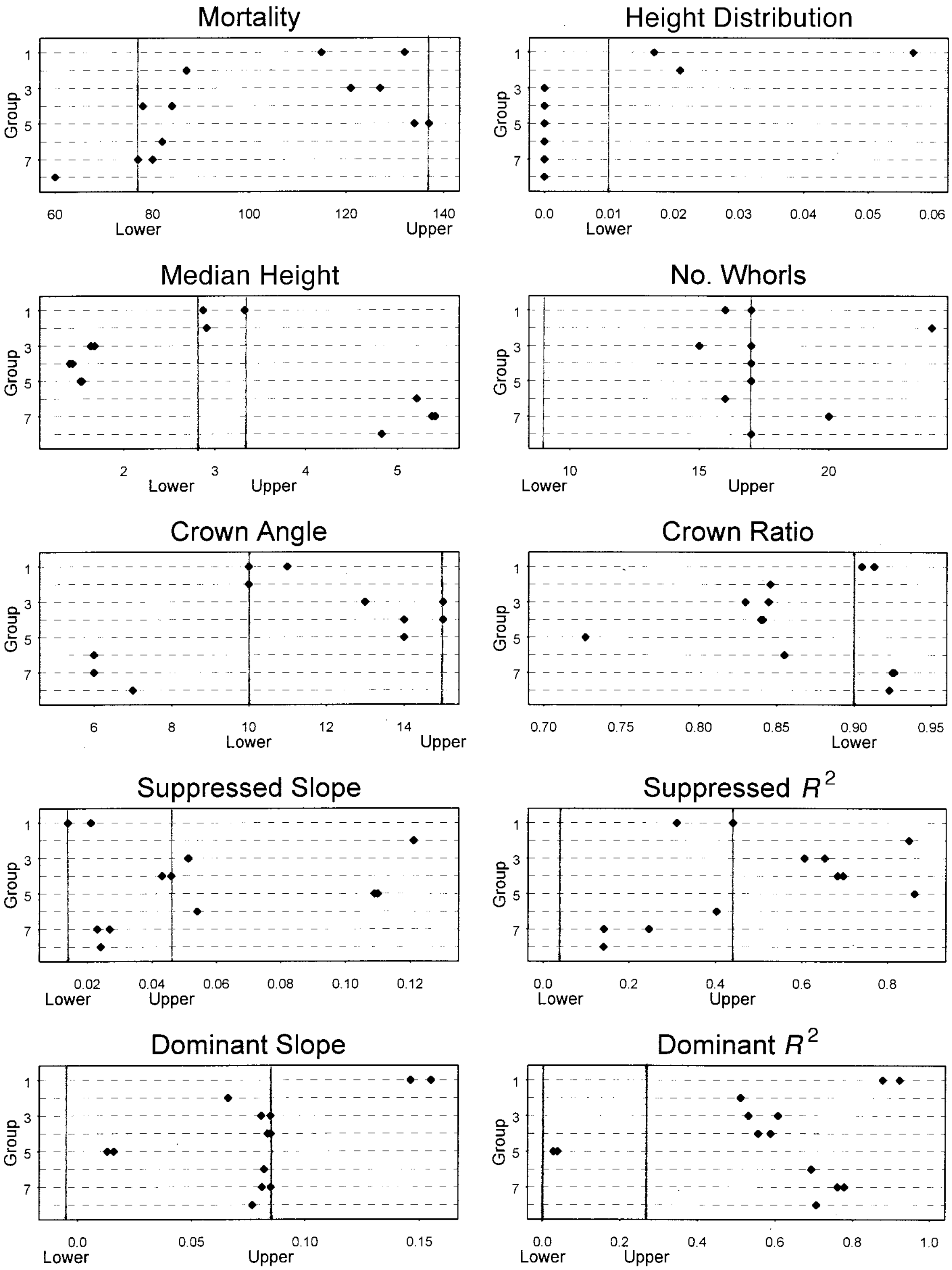

FIG. 7. Lower and upper limits of simulation results in each of the parameterization groups in WHORL2's Pareto Optimal Set (Table 5), for each criterion. Binary error interval limits (Table 2) are marked by vertical lines labeled "Lower" and "Upper" on the *x*-axis of each plot. Limits beyond the graph range are not shown. Example: Group 8's simulation underpredicts mortality and overpredicts median live height.

TABLE 6. Effect on WHORL2's Pareto Optimal Set of reducing the binary error intervals of specific criteria.

Criterion error interval reduction	Change in Pareto Optimal Set
Increase mortality lower bound	Group 8 > Groups 6 and 7
Increase crown ratio lower bound	Group 6 > Groups 7 and 8
Decrease no. live whorls upper bound	Group 7 > Groups 4 and 8
Decrease suppressed slope upper bound	Group 5 > Groups 4 and 6
Decrease dominant slope upper bound	Group 1 > Groups 2, 3, 4, 6, 7, 8

Note: ''Group A > Group B'' means that the parameterizations in Group A dominate the parameterizations in Group B under the interval reduction proposed in the left column.

any collection of the general stand and open-grown criteria (the first eight criteria). The key feature in application of POMAC is selecting a set of informative and well-formulated criteria. These will be determined partly by the context of the model's application and partly by current understanding of the modeled phenomenon. In this sense, individual-based models are extraordinarily amenable to assessment using POMAC as their many outputs provide a wealth of criteria. Criteria selection and formulation decisions are themselves assessed in the cycle (Fig. 2) and can be revised. For example, more data will be gathered to reformulate WHORL2's growth rate criteria and interval limits.

The insights obtained from using POMAC are due to considering a number of criteria simultaneously. Reducing the optimization problem by disaggregating the criterion set, or working on clusters of criteria sequentially, would undermine the very purpose of calculating the Pareto Optimal Set. For this reason, the only requirements for POMAC are computational facilities sufficient to generate the Pareto Optimal Set. While WHORL(2)'s moderate size and well documented process structure (Sorrensen-Cothern et al. 1993) aided assessment, they are not requirements of the technique. For larger models composed of more independently functioning components than those in WHORL, it may be best first to apply POMAC to the individual components. Each assessment should utilize criteria specifically focusing on the adequacy of that component's dynamics. Once each component is deemed ''adequate,'' a more comprehensive set of criteria should be chosen for applying POMAC to the complete model structure.

The Pareto Optimal Set can also be used as a tool for comparing model structures. In model construction, especially with large models, one is often faced with choosing from competing mathematical representations: e.g., should a simpler representation of photosynthesis be used in a model of forest productivity or a more complex one, containing more parameters for estimation and having greater computational demand? Such decisions of adequate representation tend to be overshadowed by a focus on sensitivity analysis and whether the parameters of a given formulation can be inferred from the available data, perhaps as a result of the tendency for single criterion model assessment. A more direct assessment of the representations is to construct models with each formulation, generate their Pareto Optimal Sets with respect to the same criterion set, and then compare the Pareto Optimal Sets. Are both structures able to satisfy the same criteria in the same ways?

Similarly, the Pareto Optimal Set can be used to compare the capabilities of wholly different model structures in satisfying a common set of criteria. For example, in model aggregation (constructing a simplified model that adequately captures a more complex model's dynamics) characteristics considered essential can be selected as the criteria and the simplified model revised until it simultaneously satisfies the complete set.

ACKNOWLEDGMENTS

The first author is extremely indebted to Mark Kot's initial mention of Pareto Optimality, as well as the University of Washington's Center for Quantitative Science in Forestry, Fisheries, and Wildlife for providing T.A. support and the Interdisciplinary Graduate Program in Quantitative Ecology and Resource Management for scholarship support while this research was being undertaken; it was otherwise unfunded. The authors would also like to thank Marianne Turley for her close reading and detailed editorial comments.

The first author dedicates this paper to the memory of Bill Heintzelman.

LITERATURE CITED

Ågren, G., R. McMurtrie, W. Parton, J. Pastor, and H. Shugart. 1991. State-of-the-art of models of production–decomposition linkages in conifer and grassland ecosystems. Ecological Applications **1**:118–138.

Bäck, T., F. Hoffmeister, and H. P. Schwefel. 1991. A survey of evolution strategies. Pages 2–9 *in* R. Belew and L. Booker, editors. Proceedings of the Fourth International Conference on Genetic Algorithms. Morgan Kaufmann, Los Altos, California, USA.

Beck, M. B. 1985. Lake eutrophication: identification of tributary nutrient loading and sediment resuspension dynamics. Applied Mathematics and Computation **17**:433–58.

———. 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research **23**:1393–1442.

Beck, M. B., and E. Halfon. 1991. Uncertainty, identifiability and the propagation of prediction errors: a case study of Lake Ontario. Journal of Forecasting **10**:135–161.

Cleveland, W. S. 1979. Robust locally-weighted regression and smoothing scatterplots. Journal of the American Statistical Association **74**:829–836.

DeAngelis, D. L., and L. J. Gross, editors. 1992. Individual-based models and approaches in ecology. Chapman and Hall, New York, New York, USA.

DeAngelis, D. L., and K. A. Rose. 1992. Which individual-based approach is most appropriate for a given problem? Pages 67–87 *in* D. L. DeAngelis and L. J. Gross. Individual-based models and approaches in ecology. Chapman and Hall, New York, New York, USA.

Fitz, H., E. DeBellevue, R. Costanza, R. Boumans, T. Maxwell, L. Wainger, and F. Sklar. 1996. Development of a general ecosystem model for a range of scales and ecosystems. Ecological Modelling **88**(1–3):263–295.

Fogel, D. B. 1994. Applying evolutionary programming to selected control problems. Computers and Mathematics with Applications **27**(11):89–104.

Ford, E. D., and K. A. Sorrensen. 1992. Theory and models of inter-plant competition as a spatial process. Pages 363–407 *in* D. L. DeAngelis and L. J. Gross. Individual-based models and approaches in ecology. Chapman and Hall, New York, New York, USA.

Gentil, S., and G. Blake. 1981. Validation of complex ecosystem models. Ecological Modelling **14**:21–38.

Goldberg, D. E. 1989. Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, New York, New York, USA.

Gross, L. J., K. A. Rose, E. Rykiel, W. Van Winkle, and E. E. Werner. 1992. Individual-based modeling: summary of a workshop. Pages 511–522 *in* D. L. DeAngelis and L. J. Gross. Individual-based models and approaches in ecology. Chapman and Hall, New York, New York, USA.

Halfon, E. 1979. On the parameter structure of a large-scale ecological model. Pages 279–93 *in* G. P. Patil and M. L. Rosenzweig, editors. Contemporary quantitative ecology and related econometrics. International Co-operative Publishing, Fairland, Maryland, USA.

Holland, J. H. 1992. Genetic algorithms. Scientific American, July:66–72.

Hornberger, G. M., and B. J. Cosby. 1985. Selection of parameter values in environmental models using sparse data: a case study. Applied Mathematics and Computation **17**: 335–355.

Hornberger, G. M., and R. C. Spear. 1981. An approach to the preliminary analysis of environmental systems. Journal of Environmental Management **12**:7–18.

Jaffe, P. R., C. Paniconi, and E. F. Wood. 1987. Model calibration based on random environmental fluctuations. Journal of Environmental Engineering **114**(5):1136–1145.

Keesman, K., and G. van Straten. 1989. Identification and prediction propagation of uncertainty in models with bounded noise. International Journal of Control **49**(6): 2259–69.

Kursawe, F. 1991. A variant of evolution strategies for vector optimization. Pages 193–197 *in* Lecture notes in computer science 496: parallel problem solving from nature. G. Goos and J. Hartmanis, editors. Springer-Verlag, Berlin, Germany.

Landsberg, J. J., M. R. Kaufmann, D. Binkley, J. Isebrands, and P. G. Jarvis. 1991. Evaluating progress toward closed forest models based on fluxes of carbon, water and nutrients. Tree Physiology **9**:1–15.

Mäkelä, A. 1988. Performance analysis of a process-based stand growth model using Monte Carlo techniques. Scandinavian Journal of Forest Research **3**:315–331.

McMurtrie, R., and H. Comins. 1996. The temporal response of forest ecosystems to doubled atmospheric CO_2 concentration. Global Change Biology **2**(1):49–57.

Michalewicz, Z. 1992. Genetic algorithms + data structures = evolution programs. Springer-Verlag, Berlin, Germany.

Michalewicz, Z., C. Z. Janikow, and J. B. Krawczyk. 1992. A modified genetic algorithm for optimal control problems. Computers and Mathematics with Applications **23**(12):83–94.

Murdoch, W. W., E. McCauley, R. M. Nisbet, S. C. Gurney, and A. M. de Roos. 1992. Individual-based models: combining testability and generality. Pages 18–35 *in* D. L. DeAngelis and L. J. Gross. Individual-based models and approaches in ecology. Chapman and Hall, New York, New York, USA.

Olenik, S., and Y. Haimes. 1979. A hierarchical multiobjective framework for water resources planning. Institute of Electrical and Electronic Engineers Transactions on Systems, Man, and Cybernetics **9**:534–544.

Oreskes, N., K. Shrader-Frechette, K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science **263**(4 February):641–646.

Pacala, S. and D. Deutschman. 1995. Details that matter: the spatial distribution of individual trees maintains forest ecosystem function. Oikos **74**(3):357–365.

Reynolds, J. H. 1997. Multi-criteria assessment of ecological process models using pareto optimization. Dissertation. University of Washington, Interdisciplinary Program in Quantitative Ecology and Resource Management, Seattle, Washington, USA.

Rose, K. A., E. P. Smith, R. H. Gardner, A. L. Brenkert, and S. M. Bartell. 1991. Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty. Journal of Forecasting **10**:117–134.

Rykiel, E. J. Jr. 1996. Testing ecological models: the meaning of validation. Ecological Modelling **90**:229–244.

Schimel, D., VEMAP participants, and B. Braswell. 1997. Continental scale variability in ecosystem processes: models, data, and the role of disturbance. Ecological Monographs **67**:251–271.

Sievanen, R., T. E. Burk, and A. R. Ek. 1988. Parameter estimation in a photosynthesis-based growth model. Pages 345–352 *in* A. R. Ek, S. R. Shifley, and T. E. Burk, editors. Forest growth modelling and prediction. Volume 1. U.S.D.A. Forest Service General Technical Report NC-120.

Sorrensen-Cothern, K. A., E. D. Ford, and D. Sprugel. 1993. A process based model of competition for light incorporating plasticity through modular foliage and crown development. Ecological Monographs **63**:277–304.

Steuer, T. 1986. Multiple Criteria optimization: theory, computation, and application. John Wiley and Sons, New York, New York, USA.

Taylor, B., K. R. Davis, and R. North. 1975. Approaches to multiobjective planning in water resources projects. Water Resources Bulletin **11**(5):999–1008.

Uhry, J. P. 1989. Applications of simulated annealing in operations research. Pages 191–203 *in* J. P. Penot, editor. New methods in optimization and their industrial uses. Birkhauser Verlag, Basel, Switzerland.

van Straten, G., and K. Keesman. 1991. Uncertainty propagation and speculation in projective forecasts of environmental change: a lake-eutrophication example. Journal of Forecasting **10**:163–190.

Vincent, T. L. 1987. Renewable resource management. Pages 161–86 *in* W. Stadler, editor. Applications of multicriteria optimization in engineering and science. Plenum Press, New York, USA.

Vincent, T., and W. Grantham. 1981. Optimality in parametric systems. John Wiley and Sons, New York, New York, USA.

Yu, P. L. 1985. Multiple-criteria decision making. Plenum Press, New York, New York, USA.