CO-OP REDWOOD YIELD RESEARCH PROJECT
Department of Forestry and Conservation
College of Natural Resources
University of California
Berkeley, California 94720

Lee C. Wensel
Associate Professor

Bruce Krumland
Assistant Specialist

Research Note No. 11 April 30, 1979

DIAMETER DISTRIBUTION MODELS FOR COASTAL STANDS IN CALIFORNIA

by

Bruce Krumland and Lee C. Wensel

Abstract

Models are developed which can be used to generate diameter distributions by species "typical" of even aged young growth stands in the north coastal region of California. These models can be used to approximate the necessary input data required by the coastal stand simulation model when the only information available are broad stand descriptors such as site index, age, stems per acre, and species composition.

This note is divided into two parts. The first part provides a general overview of the objectives of this study and describes how the results can be utilized. Part II is a technical section which describes the basic models and analytical procedures.

CONTENTS

PART I. OVERVIEW AND USE.

I. INTRODUCTION

The tree simulation model being developed for coastal stands requires either raw inventory data or a modified stand table as input. In some management situations, it may be desirable to relax the specificity of the input requirements to obtain information on the "typical" course of development of a broad class of stands. In this category, we include "hypothetical" stands, i.e., stands which may be described by a few broad characteristics yet not having an on-the-ground counterpart.

Another primary objective of the redwood cooperative is to publish a collection of yield tables for stands "typical" of various stocking levels and species compositions managed under different treatment alternatives.

To use the tree simulation model in these situations requires some means of generating an initial stand table from broad stand characteristics. Research Note No. 8 described a means of predicting tree heights given diameter. This note describes some models for generating mixed species diameter distributions for young-growth even-aged, uncut stands from variables such as site index, age, stems per acre, and species composition.

The diameter distribution models described in this note, the height prediction models (Research Notes 8 and 12), and the crown estimation models (draft in progress) have all been coded into a stand generation computer model. This computer model will be made available to potential users and it will also be used as a standard for supplying initial stand descriptions to the coastal stand simulation model for general yield table construction.

II. OBJECTIVES AND USE OF THE DIAMETER DISTRIBUTION MODELS

The models presented in this report have their primary use in translating a stand description based on broad characteristics into a stand table (numbers of trees by DBH and species class) which is a portion of the necessary input data required for the coastal tree growth model. Stands generated on this basis have their primary use in situations such as

(a) Establishing regeneration stocking levels
(b) Establishing guidelines for precommercial thinning intensities
(c) Setting general standards for residual stocking levels after partial harvests

In situations where stands are not necessarily even aged, have experienced past harvesting, or where more refined stand specific estimates are desired, actual plot records will be required.

A. Overview of the Models

Two models have been developed to generate diameter distributions. The basic difference is that one uses a more refined stand description.

a) Model I - This model requires the user to specify 50 year site index, average breast high age of site trees, and stems per acre for any combination of the following species groups.

1) redwood
2) other conifers (mostly Douglas fir)
3) tanoak (includes other hardwoods except alder)
4) alder

A two stage process is then followed: (1) the specified stand information is then used to estimate the quadratic mean DBH of each species group individually unless the species group is absent (i.e., the number of stems per acre for the species group is not given). (2) the stand information plus the predicted mean DBH's are then used to estimate shape and location coefficients for a diameter distribution model for each species group.

b) Model II - This model requires everything Model I does, with the addition of a specified quadratic mean DBH for each species group if it is present. This refinement essentially reduces some error that comes from 'centering' the distribution and makes the model more stand specific.

It is emphasized that stems per acre includes all trees 4.5 feet in total height and taller.

B. Usage and Data Requirements

In order to make this model easy to use, several simplifications have been developed.

a) Stems Per Acre - Either stems per acre by species group or percentages of total stems per acre are acceptable. Using basal area or stems/acre greater than some minimum DBH can also be incorporated with the aid of tables described in Section III.

b) Site Index - Frequently, it will be desirable to express a site index value for a single conifer species only. If site indices of some species are not supplied, they will be estimated by procedures described in Appendix II.

c) Breast High Age - Age estimates based on the supplied age of at least one conifer species are also accomplished by procedures described in Appendix II.

III. USE OF BASAL AREA AND STEMS/ACRE BY STAND FRACTION

Many potential users are accustomed to thinking in terms of basal area rather than stems/acre (particularly for larger stands) or in terms of stand components greater than some arbitrary minimum DBH. Tables A. and B. have been prepared so that approximate conversions can be made.

Two of the input variables, site index and age, are used to estimate total height of dominant trees. Total height, then, is what is actually utilized in the prediction equations. These tables were prepared by supplying a top height and total number of stems per acre in several possible combinations for redwood and Douglas fir separately. The model previously described (Model I) was then used to generate a stand table consisting of stems per acre by one inch diameter classes. Summaries were then developed showing stems and basal area for the portion of the stand greater than 0, 5, and 11 inches DBH.

For example, assume that you wish to generate a stand table for a pure Douglas fir stand with a dominant height of 75 feet and 100 square feet of basal area in stems 11 inches DBH and greater. Turning to Table B for Douglas fir, we look for the entry "100" under the 8th column (basal area of stems 11 inches dbh and larger) for a dominant height of 75 feet. Referring to the 4th column (stems per acre 0 inches and over) we see that this corresponds to 200 trees per acre. To generate the stand table with 100 sq. ft. of basal area (11" dbh+) we enter 200 trees per acre.

TABLE A:. REDWOOD - PURE STANDS

Estimated stems/acre, basal area, and average (quadratic mean) dbh by dominant height and stand fraction

dom ht	min dbh	stand 0" dbh+			stand 5" dbh+		stand 11" dbh+	
		basal area	stems acre	ave dbh	basal area	stems acre	basal area	stems acre
25	0.4	12	100	4.7	8	29	0	1
25	0.0	23	200	4.6	17	51	2	4
25	0.0	31	300	4.4	22	66	4	6
25	0.0	37	400	4.1	26	78	3	6
25	0.0	43	600	3.6	27	87	3	6
25	0.0	43	800	3.2	24	84	0	4
50	2.2	53	100	9.9	52	86	30	29
50	1.6	83	200	8.8	79	148	39	39
50	1.2	104	300	8.0	97	195	42	43
50	0.9	119	400	7.4	108	232	43	46
50	0.7	129	500	6.9	114	258	40	45
50	0.4	135	600	6.4	116	276	36	42
75	3.4	57	50	14.6	57	49	50	33
75	3.0	98	100	13.4	98	96	81	55
75	2.4	161	200	12.2	160	181	122	89
75	2.1	208	300	11.3	205	258	148	113
75	1.8	244	400	10.6	238	326	161	128
75	1.6	270	500	10.0	262	388	166	137
100	4.3	95	50	18.7	95	50	92	42
100	4.0	168	100	17.6	168	99	158	77
100	3.7	231	150	16.8	231	148	213	109
100	3.5	286	200	16.2	286	196	260	138
100	3.4	335	250	15.7	335	244	301	165
100	3.2	379	300	15.2	379	290	336	191
125	5.0	129	50	21.8	129	50	127	45
125	4.6	231	100	20.6	231	100	224	85
125	4.3	320	150	19.8	320	149	307	122
125	4.1	399	200	19.1	399	198	381	158
125	3.9	470	250	18.6	470	248	445	193
125	3.8	534	300	18.1	534	296	502	225

Table E.: DOUGLAS FIR - PURE STANDS

Estimated stems/acre, basal area, and average (quadratic mean) dbh by dominant height and stand fraction

dom ht	min dbh	stand 0" dbh+			stand 5" dbh+		stand 11" dbh+	
		basal area	stems acre	ave dbh	basal area	stems acre	basal area	stems acre
25	0.7	8	100	4.0	4	20	0	0
25	0.2	11	200	3.2	3	18	0	1
25	0.0	15	300	3.1	3	22	0	1
25	0.0	18	400	2.9	2	18	0	2
25	0.0	20	600	2.5	0	3	0	3
25	0.0	19	800	2.1	0	4	0	4
50	1.9	37	100	8.3	36	74	15	16
50	1.4	59	200	7.4	54	131	16	20
50	1.2	74	300	6.8	65	176	13	20
50	1.0	84	400	6.2	71	210	8	16
50	0.9	90	500	5.8	72	232	4	12
50	0.7	93	600	5.3	70	243	0	9
75	3.4	48	50	13.3	48	48	39	26
75	3.0	84	100	12.5	84	93	65	46
75	2.5	141	200	11.4	140	178	100	78
75	2.3	185	300	10.6	181	258	119	100
75	2.1	217	400	10.0	211	330	128	114
75	2.0	240	500	9.4	231	395	127	120
100	3.7	72	50	16.3	72	49	66	34
100	3.3	129	100	15.4	129	97	114	64
100	3.0	177	150	14.7	177	144	154	92
100	2.8	219	200	14.2	219	190	187	116
100	2.7	256	250	13.7	255	236	214	138
100	2.6	289	300	13.3	287	281	236	158
125	4.2	106	50	19.7	106	50	102	40
125	3.8	191	100	18.7	191	99	182	78
125	3.5	265	150	18.0	265	148	249	113
125	3.3	331	200	17.4	331	196	308	146
125	3.2	388	250	16.9	388	244	359	178
125	3.1	439	300	16.4	439	293	402	207

IV. SOME ILLUSTRATIVE RESULTS

As an example of the kinds of results that are available with this model, suppose we specified a stand with 600 stems/acre with a redwood site index of 100 feet and a breast high age of 15 years. Let us further suppose that one third of the trees are redwood, on third are Douglas fir, and one third are tanoak.

In this case, the computer models would automatically make the following estimates.

	Site Index	Breast High Age
Douglas Fir	127	12
Tanoak	70	13

Next, by using the diameter distribution models described in Part II., the following stand table is generated.

Stems Per Acre

2 inch DBH class	Redwood	Douglas Fir	Tanoak
0 - 2	45	40	32
2 - 4	51	86	106
4 - 6	40	58	56
6 - 8	28	16	9
8 - 10	18		
10 - 12	11		
12 - 14	6		
14 - 16	2		

Figure 1. shows the smoothed diameter distributions of each of the three species groups separately.

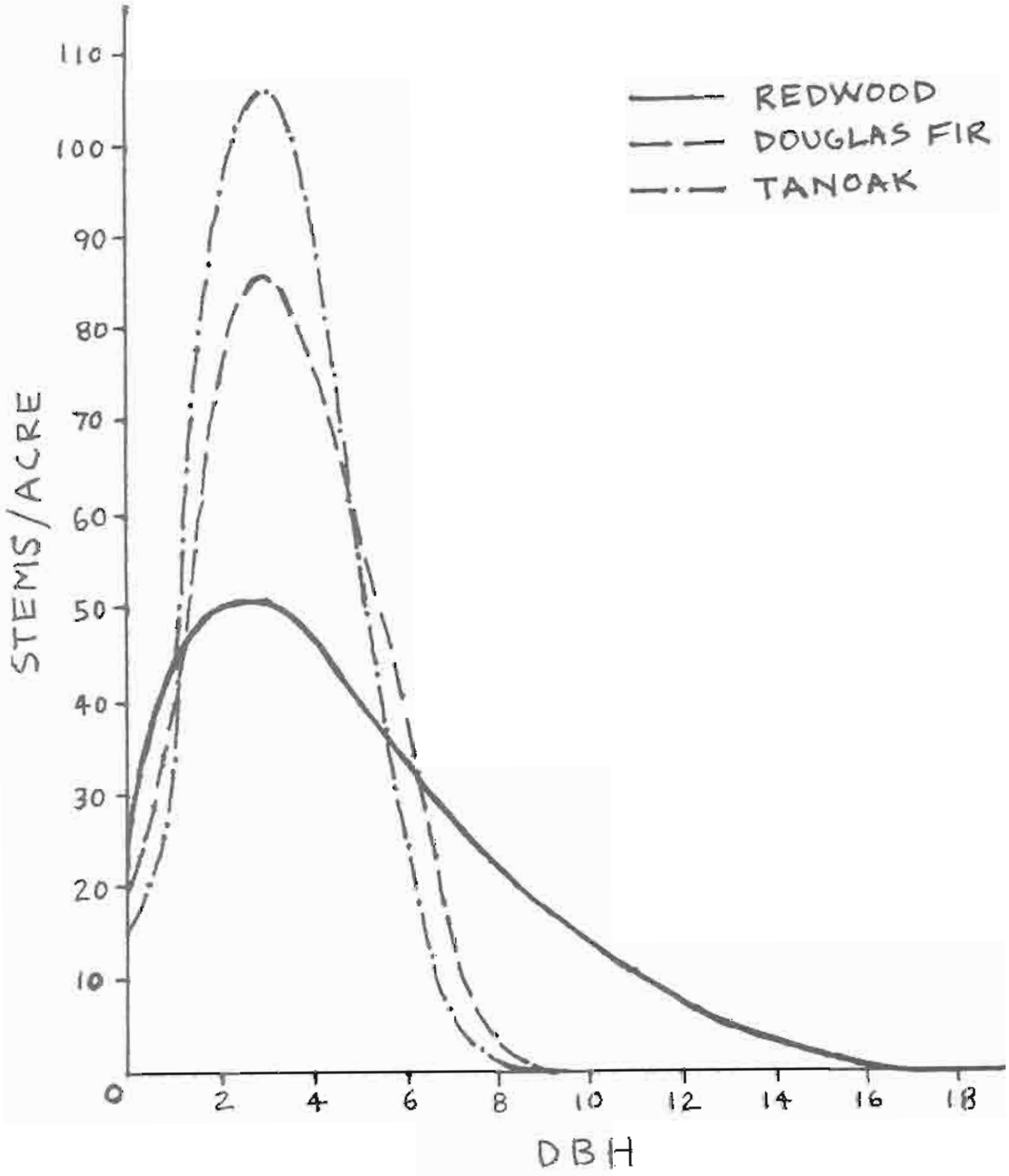

Figure 1: Estimated diameter distribution for a redwood site index of 100 feet and a breast high age of dominant redwoods of 15 years. Two hundred stems per acre were specified for each of the species groups: redwood, Douglas fir, and tanoak.

PART II. MATHEMATICAL MODEL DEVELOPMENT

I. THE DIAMETER DISTRIBUTION MODEL

The literature in forest mensuration contains many models which are proposed as being useful in describing diameter distributions of forest stands (Meyer 1930, Nelson 1964, Schnur 1934, Bliss and Reinker 1964, Clutter and Bennet 1965, Bailey and Dell 1973, Hafley and Schreuder 1977). After some preliminary analysis, the Weibull probability density function was chosen as the base model in this study because of its simplicity, flexibility, and generally satisfactory results. In its basic form, this model can be written as

$$F(d) = 1. - \exp(-((d-a)/b)^{c}) \tag{1}$$

where

F(d) = Percentage of trees in a forest stand less than or equal to 'd' inches DBH

a,b,c = Parameters to be estimated

exp(x) = 2.17828 raised to the power of 'x'

The parameter 'a' is the smallest diameter in the stand. The parameter 'b' is a measure of central location such that approximately 63% of the trees are less than 'b' inches in diameter. The parameter 'c' controls the shape of the diameter distribution. For a value of 'c' equal to 3.6, the diameter distribution approximates the normal probability distribution in shape. For values of 'c' less than 1, the curve is inverse J-shaped. For values of 'c' between 1 and 3.6, the distribution is unimodal and skewed to the right. As 'c' becomes greater than 3.6, the curve is unimodal and progressively skewed to the left.

If, for example, we wish to generate a stand table from this function for a stand with 'N' trees per acre, the number of stems 'N_d' between any two diameters d_1 and d_2 ($d_2 > d_1$) is

$$N_d = N(F(d_2) - F(d_1)) \tag{2}$$

$$= N(\exp(-((d_2 - a)/b)^{c} - \exp(-((d_1 - a)/b)^{c})$$

By incrementally altering d_1 and d_2, we can essentially create a stand table.

II. OVERVIEW OF MODEL DEVELOPMENT

The procedures utilized in developing a diameter distribution model for coastal stands consisted of the the following three steps:

1) Selecting plot records thought to be typical of young growth even-aged stands in the region from available growth plot records.

2) Estimating the parameters in the Weibull probability density function for each species group in each plot.

3) Using the collection of parameter estimates from step 2 as observations for additional models relating Weibull parameters to common stand attributes such as site index, age, stems per acre, and species composition.

III. PLOT SELECTION AND DATA SOURCES

Initial screening produced 583 plot records for subsequent analysis. These records were single measurements from growth plots maintained by cooperators in the Redwood Yield Research Project. Approximately half of the plots were located in Humboldt and Del Norte counties and the other half were from Mendocino. Initial criteria were that (1) average breast-high age and 50-year base age site index estimates were available for at least one conifer species group on candidate plots; (2) no evidence that past harvesting had taken place since regeneration; 3) no old growth trees were standing on the plots.

Next, species groups within each plot were graphically screened with the aid of an interactive computer plotting routine. Based on data limitations, the similarity of diameter distributions of several species, and the low occurence of several species on many of the plots, the following four species groups were recognized:

1) redwood
2) Douglas fir (includes other conifers)
3) tan oak (includes other hardwoods except alder)
4) alder

Screening was designed to identify the species groups on each plot for which diameter distribution parameters could be adequately estimated. Species groups on each plot were not considered for modeling if (1) there were an insufficient number of stems to define a distribution or (2) arbitrary selection of a lower diameter limit for field measurements resulted in truncating the diameter distribution to the right of the mode. This screening process left the following number of diameter distribution observation sets by species group:

redwood	213
Douglas fir	192
alder	25
tan oak	80

IV. PLOT SPECIES GROUP PARAMETER ESTIMATION

Parameter estimates of each species group on each plot selected for modeling were estimated by an iterative nonlinear least squares procedure fitted to the empirical cumulative plot diameter distribution. Special procedures were required in cases where field measurement specifications resulted in truncating the diameter distribution to a diameter limit greater than the "natural" minimum diameter. The resulting parameter estimates were then used in estimating quadratic mean diameter, basal area and stems/acre. These procedures are described in Appendix I.

For species groups on selected plots whose distribution parameters were not estimated, basal area and stems per acre below the minimum recorded diameter were estimated by graphical analysis. These estimates plus actual recorded measurements were then used to obtain an estimate of total per acre basal area and stems per acre.

Dominant height (height predicted by site index equations), average breast-high age of dominants, and fifty-year breast high age base site index were also summarized for each plot species group. These descriptors were estimated from actual plot measurements or by the conversions described in Appendix II.

V. GENERAL STAND DIAMETER DISTRIBUTION MODEL

The final phase of this study consisted of developing a system of models to relate parameters of the Weibull plot distributions to broad stand characteristics. Potentially, the system of models would require 12 prediction equations (3 parameters x 4 species groups). Several items were given explicit consideration during construction.

1) It was initially assumed that the diameter distribution of any species group, whether in a monoculture or a mixture, could adequately be approximated by a Weibull density function. Based on visual examinations of plots, this assumption appeared to be reasonable.

2) The general models would use stand descriptors that most forest managers are familiar with. This limitation was necessary from a practicable standpoint.

3) Lastly, it was recognized that there would be interactions among species in mixed stands. The relative size and abundance of other species were presumed to influence the shape and location of a single species diameter distribution. Preliminary analysis supported this tenet and also indicated that the parameters of species diameter distributions on individual plots were themselves correlated. Incorporating these correlations in a multivariate framework would be desirable from the standpoint of efficiency in model development. However, as all species were not present on all plots and some species present on some plots did not have distribution parameter estimates, some "less than optimal" estimation procedures were resorted to.

Attempts to estimate distribution parameters directly met with little success. However, it was found that quadratic mean DBH of individual species could be estimated fairly well as a function of the dominant species height, stems per acre and average dominant height of all species combined. Next, it was found that while distribution parameters themselves could not be estimated with an acceptable level of precision, predicted diameters at fixed percentage points could be estimated with a fair degree of success. Using the estimated values of a, b, and c for a plot species group and manipulating equation (1) gives the following relationship:

$$d_i = a+b(-\ln(1-p_i))^{1/c} \tag{3}$$

where

d_i = predicted diameter at a percentage point p_i

$\ln(1-p_i)$ = natural logarithm of $(1-p_i)$

Hence, rather than estimate the three species parameters directly, prediction equations for three percentiles could be used to solve for parameter estimates of a, b, and c. Lohrey and Bailey (1977) used a similar approach with good results. Solutions are greatly simplified by a judicious choice of the three percentage points. The points used in this study are:

d_i	Percentage Point (p_i)
d_1	.89000
d_2	.60302
d_3	.32070

The reason for this choice is detailed in Appendix III. Given the triplet of predictions for a given species ($\hat{d}1$, $\hat{d}2$, $\hat{d}3$), closed form expressions of the parameter estimates are as follows:

$$\hat{c} = .870932/\ln((\hat{d}_1-\hat{d}_2)/(\hat{d}_2-\hat{d}_3)) \tag{4}$$

$$\hat{b} = (\hat{d}_1-\hat{d}_2)/(2.20727^{1./\hat{c}} - 2.20727^{-.1/\hat{c}}) \tag{5}$$

$$\hat{a} = \hat{d}_1-\hat{b}(2.20727^{1/\hat{c}}) \tag{6}$$

A system of equations for expressing the diameter distribution model was subsequently developed. The broad stand descriptors comprising the basic independent variables for the model are site index, breast high age, and stems/acre by species.

Not surprisingly, it was found that diameters at fixed percentage points were correlated most with quadratic mean diameter. Consequently, the following two-stage procedure was utilized in developing the equation system for the model.

1. Estimate quadratic mean DBH (D) for each species group solely as a function of independent variables

2. Estimate diameters at fixed percentage for each species as a function of predicted quadratic mean diameters plus other independent variables

A. Quadratic Mean DBH Model.

After some experimentation, the following functional form based on a general sigmoidal relationship was found to give reliable and logical estimates.

$$D_s = a_1(H_s)^{a_2} \{1 + \exp(a_3 N + a_4 H_s/AH + a_5)\}^{a_6} \qquad (7)$$

where

D_s = quadratic mean diameter of species 's'

N = total stems/acre

H_s = dominant height of species 's' predicted from site curves

AH = average dominant height of all species weighted by stems per acre

$$= \sum_{s=1}^{4} H_s N_s / N \qquad (8)$$

N_s = stems/acre of species 's'

a_i = species dependent coefficients

Parameter estimates and a statistical summary for each of the four species groups are shown in Table 1. More complicated models involving stems/acre and dominant height of each species group separately resulted in insignificant reductions in the mean square error.

Table 1

Parameter estimates and a statistical summary by species
for the quadratic mean diameter model

Species	a_1	a_2	a_3	a_4	a_5	a_6	R^2	$S_{y.x}$	Sample Size
Redwood	.488	.93	.60	-.83	-1.10	4.23	.62	3.6	214
Douglas Fir	.470	1.00	.43	-.12	-1.34	4.35	.71	3.9	192
Tanoak	.689	.87	.28	.40	-1.48	2.60	.40	2.8	80
Alder	.446	.98	.25	.17	-1.49	3.36	.63	2.3	25

B) Diameter Percentile Estimators (Model I)

As noted earlier, percentile diameters were most highly correlated with quadratic mean diameter of individual species. Based on some extensive screening processes, the following equation forms were used for each of the four species groups:

$$d_{is} = b_1\hat{D}_s + b_2 \ln(N) + b_3 \ln(A_s) \qquad (9)$$

where

d_{is} = diameter at percentage point "i" for species "s"

$\hat{D}_s$ = predicted quadratic mean DBH of species "s"

N = total stems per acre

A_s = breast high age of species "s"

$\ln(x)$ = natural logarithm of "x"

b_i = species dependent parameters to be estimated

The parameter b_3 for the hardwood groups was found to be generally insignificant after the inclusion of D_s and N, so it was dropped from the prediction equations for these groups.

C) Estimation Procedures

As noted earlier, it was not feasible to estimate the system of twelve equations simultaneously. However, it was possible to simultaneously estimate the parameters for the three equations comprising a single species group. The basic reasons for doing this are:

(a) Variances of parameter estimates are smaller than in the case where the parameters were estimated independently.

(b) Conventional least squares applied to this three equation system would minimize the sums of squared residuals of each equation separately. Simultaneous estimation minimizes the sums of squared residuals of all three equations at the same time (with weights being assigned to the observations inversely proportional to their variance - covariance terms).

(c) The resulting differences between parameter estimates in each of the three equations are more consistent than in the case where the parameters in the three equations are estimated separately.

The procedure used in estimation is sometimes called generalized least squares applied to systems of equations. The details are somewhat tedious and are not described here. Readers desiring more information are referred to Maddala (1977, pp. 465-467). Parameter estimates and a statistical summary are shown in Table 2.

4) Constraints on Predictions

The systems of models previously described have been coded into an interactive computer stand generator that allows the user to rapidly examine simulated diameter distributions under different combinations of stand descriptors. In some situations (stands 15 years or less at high stems/acre levels) the estimated value of the parameter 'a' is sometimes negative. When this happens, some alternate estimation procedures are used. These procedures are described in Appendix IV.

Table 2

Parameter estimates and statistical summary by species for the DBH percentile models using predicted quadratic mean DBH as an independent variable (Model I)

Redwood	b_1	b_2	b_3	R^2	$S_{y.x}$ (inches)	Sample Size
d_1	.773	-1.439	4.588	.65	4.8	214
d_2	.470	-1.711	4.645	.73	3.0	214
d_3	.304	-1.701	4.336	.70	2.7	214
			Overall	.68	3.6	642
Douglas Fir						
d_1	.826	-1.409	4.345	.77	4.5	192
d_2	.689	-.923	2.698	.79	3.2	192
d_3	.546	-.721	2.005	.75	2.7	192
			Overall	.77	3.6	576
Tanoak						
d_1	1.413	.076	1/	.25	4.2	80
d_2	1.100	.024	1/	.26	3.3	80
d_3	.869	.009	1/	.27	2.7	80
			Overall	.26	3.5	240
Alder						
d_1	1.051	.262	1/	.69	1.9	25
d_2	.934	.073	1/	.71	1.6	25
d_3	.852	-.082	1/	.69	1.5	25
			Overall	.70	1.7	75

1/ Tanoak and alder equations were estimated without the b_3 parameter.

VI. USE OF BASAL AREA FOR REFINEMENTS (Model II)

It was originally anticipated that two general models would be developed; one using stems per acre as a density term and the other using basal area. In the use of stems per acre as a density term, it seems fairly logical that the more stems that are present, the smaller is the average diameter (everything else being constant). Analysis indicated, however, that the relationship between basal area and average diameter is not necessarily single valued. In other words, for a given level of basal area, stands may exist with a small number of large trees or a large number of small trees, the former case having a larger average diameter than the latter. This basic indeterminancy precluded the development of a diameter distribution model based on basal area.

If basal area and stems per acre are known (or given), some efficiency can be gained because quadratic mean DBH can be derived directly and the first stage model which is used to estimate mean DBH can be bypassed. Another system of diameter percentiles was estimated using actual rather than predicted quadratic mean DBH (Model II). These models have the same form and were estimated by the same procedures as the previous ones. The coefficients are shown in Table 3.

Some basic relationships between stems/acre and basal area are tabled in Part I. The procedure used in computing distribution parameters is slightly different under Model II. The parameters 'b' and 'c' are calculated as under Model I. The parameter 'a', however, is computed as a function of the given quadratic mean diameter and estimates of 'b' and 'c' (see Appendix IV). This procedure ensures compatibility with parameter estimates and specified quadratic mean DBH.

VII. VALIDATION

An attempt to validate this model presents some difficulties for several reasons:

a) The models presented here will be used to generate an initial stand description from which forecasts of future yields will be made. As what is "ultimately" in question is the effect of the initial description on properties of future yield estimates, it is unclear at this time how such a hypothesis might be tested.

b) While not formally stated, underlying tenets of this study are 1) tree diameters in coastal stands can be generated by a Weibull process and 2) the models developed here provide an adequate description of this process. As the data used in this study do not necessarily represent a "true" random sample of stands currently existing or stands that may exist in the future, tests of goodness of fit cannot be thought of as providing "complete" proof of model adequacy. Nonetheless, several tests were made to evaluate model performance.

A. Graphical Analysis

Histograms of actual plot diameter frequency distributions were compared with those predicted by the models. There was in general a

Table 3

Parameter estimates and statistical summary by species for the DBH percentile models using actual quadratic mean DBH as an independent variable (Model II)

	b_1	b_2	b_3	R^2	$S_{y.x}$	Sample Size
Redwood						
d_1	1.426	.451	-.933	.97	1.4	214
d_2	.891	-.499	1.093	.95	1.3	214
d_3	.618	-.807	1.700	.85	1.9	214
			Overall	.94	1.6	642
Douglas Fir						
d_1	1.351	.075	-.104	.98	1.3	192
d_2	1.011	-.002	-.060	.99	.7	192
d_3	.769	-.077	.082	.93	1.5	192
			Overall	.97	1.2	576
Tanoak						
d_1	1.295	.045	1/	.98	.6	80
d_2	1.043	-.045	1/	.99	.3	80
d_3	.839	-.065	1/	.97	.6	80
			Overall	.98	.5	240
Alder						
d_1	1.153	.194	1/	.99	.4	25
d_2	1.018	.024	1/	.99	.2	25
d_3	.935	-.138	1/	.99	.3	25
			Overall	.99	.3	75

1/ Tanoak and alder equations were estimated without the b_3 parameter.

surprising degree of conformance.

B. Comparisons With Other Studies

Model I has two principal components: 1) the quadratic mean DBH model which can be thought of as "centering" the distribution and 2) the percentile models which provide the shape. Lindquist and Palley (1967) published yield tables which included average quadratic mean DBH by age and site. While the authors of this report do not necessarily agree with the implications of yield tables concerning growth, they do feel that yield tables provide a reasonable description of stand attributes based on the underlying sample data.

A comparison was made for redwood to see how well the average DBH's predicted in this study compared with those in yield tables. Table 4 shows the results of this comparison. In general, there is quite close agreement although this studies' prediction are slightly lower.

Table 4.

Comparisons of quadratic mean DBH of all stems 4.5" DBH and greater predicted by this study (Model I) for pure redwood stands with Lindquist and Palley (1967) empirical yield table estimates by site and breast height age.[1/]

	Lindquist and Palley Site Index			
B.H. Age	140	160	180	200
	(Yield Table/Prediction)			
20	9.3/8.7	10.1/9.4	10.8/9.8	11.5/10.2
30	12.4/11.7	13.7/12.6	14.8/13.5	15.8/14.2
40	14.6/13.8	16.2/15.0	17.6/15.9	18.8/16.8
50	16.4/15.7	18.2/17.0	19.7/18.0	21.0/19.0
60	18.0/17.4	19.9/18.7	21.6/19.9	23.0/19.2
70	19.4/18.8	21.4/20.2	23.1/21.4	24.5/22.6

[1/] As the coefficients and predictions for Model I were based on all stems, the procedure used for quadratic mean DBH of stems 4.5" DBH and greater was as follows:

1) Convert Lindquist and Palley site indices to 50 year base age indices by conversions described in Research Note No. 5 (Krumland and Wensel 1977).

2) Use the equation system to estimate stems/acre greater than or equal to 4.5" DBH given a total number of stems/acre.

3) Vary the total stems per acre until the estimated number of stems/acre greater than 4.5" were equal to yield table estimates.

4) Compute quadratic mean DBH by numerical means based only on the numbers of trees 4.5" and greater.

C. Tests of "Goodness of Fit"

Further tests were made to see how the distributions predicted by these models compared with actual empirical plot distributions (the empirical distributions refer to the empirical frequency distributions of each species on each plot, not the individual plot Weibull functions fitted to this data). For Model I, there are three potential (not necessarily independent) sources of error: 1) the estimated distributions are not centered properly, 2) the estimated distributions have the wrong shape, 3) the empirical distributions could not conceivably have been generated by a Weibull process. Model II, which uses quadratic mean DBH as supplied as an independent variable, largely has items 2 and 3 listed above as the principal sources of error.

The test used is based on the numbers of trees used to make the empirical frequency distribution and the maximum absolute difference between empirical and predicted cumulative frequency. This difference is called the Kolmogorov - Smirnoff statistic and tables have been prepared (see Hogg and Tanis, 1977) showing the probability that the empirical distribution is different from predicted ones at various significance levels based on this statistic. Table 5 shows the results of this test for Redwood and Douglas fir and each of the two models. Considering that several plots had fewer than 10 trees and a very loose selection criterion was employed for including plots in this study, these results are considered to be quite acceptable.

Table 5

Percentage of estimated distributions not significantly different from empirical ones by model, species, and significance level.

	Model I		Model II	
	Significance Level			
	.05	.01	.05	.01
Redwood	68%	86%	77%	90%
Douglas Fir	83%	95%	77%	91%
TOTALS	73%	89%	77%	90%

Literature Cited

Bailey, Robert L., and T.R. Dell
1973. Quantifying diameter distributions with the Weibull function. Forest Sci. 19(2):97-104.

Bliss, C.I., and K.A. Reinkler
1964. A log-normal approach to diameter distributions in even-aged stands. Forest Sci 10:350-360.

Clutter, J.I., and F.A. Bennett
1965. Diameter distributions in old-field slash pine plantations. Ga For Res Counc Rep 13, 9p.

Curtis, Robert O. et al.
1973. Which dependent variable in site index-height-age regressions? Forest Sci. 20:74-87.

Ek, Alan R. et al.
1975. Solving for Weibull diameter distribution parameters to obtain specified mean diameters. Forest Sci. 21:290-292.

Hogg, Robert V., and Elliot A. Tanis
1977. Probability and statistical inference. Macmillan Publishing Company, New York, New York. 450p.

Krumland, Bruce K., and Lee C. Wensel
1977. Height growth patterns and fifty year base age site index curves for young growth coastal redwood. Res. Note No. 4.

Krumland, Bruce and Lee C. Wensel
1977. Procedures for estimating redwood and Douglas fir site indexes in the north coastal region of California. Res. Note No. 5.

Lindquist, J., and M. N. Palley
1967. Prediction of stand growth of young redwood. Bull. Calif. Exp. Sta., No. 831, p64.

Lohrey, Richard E., and Robert L. Bailey
1977. Yield tables and stand structure for unthinned longleaf pine plantations in Louisiana and Texas. U.S.D.A. For. Ser. Res. Paper SO-133. Southern Forest Experiment Station. 12p.

Porter, D.R., and H.V. Wiant, Jr.
1965. Site index equations for tanoak, pacific madrone and red alder in the redwood region of Humboldt County, Calif. J. For. 63(4):286-287.

Maddala, G. S.
1977. Econometrics. McGraw-Hill. 516p.

Meyer, H.A.
1952. Structure, growth, and drain in balanced uneven-aged forests. J. Forest 50:85-92.

Nelson, T.C.
1964. Diameter distribution and growth of loblolly pine. Forest Sci 10: 105-115.

Schnur, G.L.
1934. Diameter distributions for old-field loblolly pine stands in Maryland. J Agric Res 49:731-743.

Schreuder, H. T., and W. L. Hafley
1977. A useful bivariate distribution for describing stand structure of tree heights and diameters. Biometrics 33:471-478.

Appendix I

Procedures for Fitting Weibull Parameters to Plot Species Groups

In the Weibull density function for each plot species group, we have:

$$F(d) = a - \exp(-((d-a)/b))^{c} \qquad \text{(I)}$$

where

a,b,c = Model parameters to be fitted with 'a' being the minimum DBH

d = DBH

F(d) = Percentage of trees in the species group between 'd' and 'a' inches DBH

= N_d/N

N_d = Number of trees less than or equal to 'd' inches in the species group

N = total number of stems in the species group

In fitting this model to data, we often do not know the minimum diameter present. Hence, what can actually be derived from plot records is

M_d = Numbers of trees/acre between diameter a^* and d

where

a^* = Minimum DBH recorded, often arbitrarily set in measurement specifications

and

M = Total number of trees/acre greater than or equal to a^* in DBH

The values for M and all the M_d can essentially be determined from plot measurement. The number of trees between 'a' and a^* inches (α) is unknown. The truncated and actual distribution observation points are related by

$$\frac{M_d + \alpha}{M + \alpha} = N_d/N = 1 - \exp(-((d-a)/b)^{c})$$

Solving this expression for α gives:

$$\alpha = (M(1 - \exp(-((d-a)/b)^c\ ) - M_d)/\exp(-((d-a)/b)^c) \qquad \text{(II)}$$

However, we also know that α is equal to

$$\alpha = (M + \alpha)(1-\exp(-((a^* - a)^c) \qquad \text{(III)}$$

$$= M(1. - \exp(-((a^* - a)/b)^c))/\exp(-((a^* - a)/b)^c)$$

Setting II equal to III and solving yields

$$F_m(d) = M_d/M = 1. - \exp[b^{-c}((a^* - a)^c - (d-a)^c)] \qquad \text{(IV)}$$

Equation IV then is a modified Weibull function that can be fitted to actual truncated plot records to provide estimates of a, b, and c. We note that

$$F_m(\infty) - F_m(a^*) = 1$$

So it satisfies a basic property of distribution functions.

Equation IV was used as the functional form for estimating the 3 Weibull parameters for each plot species group. The value for α was computed from equation III using parameter estimates. An estimate of stems per acre was computed as

$$\text{estimated } N = M + \alpha$$

The quadratic mean diameter ($\bar{D}$) is the square root of the expected value of tree diameter squared.

$$\bar{D} = E(d^2)^{1/2}$$

Ek et al (1975) gives an expression of $\bar{D}$ as a function of Weibull parameters. This value was computed numerically using actual parameter estimates as

$$\bar{D} = (b^2\Gamma(1 + 2/c) + 2ab\Gamma(1 + 1/c) + a^2)^{1/2}$$

where

$$\Gamma(x) = \text{gamma function of } x$$

$$= \int_0^{\infty} e^{-t} t^{x-1} \delta t$$

Total stand per acre basal area (B) was then estimated as

$$B = .005454(M + \alpha)\bar{D}^2$$

Appendix II

Site, Height and Breast High Age Conversions

This appendix describes the procedures used to estimate missing measurements for the four principle species groups used in this study.

A. Dominant Height

Dominant height (H) is expressed as a function of fifty year breast high age site index (S) and breast high age (A). The basic model used is:

$$H = B1\{1. - (1. - (S/B1)^{B3})\exp((A - 50.)B2)\}^{1/B3} \qquad (II-1)$$

where

$$B1 = a_1 S^{a_2}$$

$$B2 = a_3 S^{a_4}$$

$$B3 = a_5 S^{a_6}$$

a_i = species dependent coefficients

This model was fitted to redwood data as described in Research Note No. 4 (Krumland and Wensel 1976). King's Douglas Fir site index equation (King 1966) was used to estimate heights for ages 10, 20 ... 100 at 10 foot site increments for site indices 80 through 140. These generated points were then fitted to obtain the coefficients for equation II-1. For alder, the same procedure was used using the equations supplied from Curtis et. al. (1974). For tan oak, the site index equations of Porter and Wiant (1965) were used after inverting and adjusting total age to breast high age. Coefficients for all four species are shown in Table II-1.

Table II-1

Coefficients by species for equation II-1

Species	Coefficients					
	a_1	a_2	a_3	a_4	a_5	a_6
Redwood	9.44	.68	-.00118	.46	.64	.15
Douglas Fir	2.22	.94	-.00167	.47	14.02	-.53
Alder	2.17	.88	-.00158	.58	7.48	-.304
Tan Oak	6.00	.84	-.00180	.32	1.63	-.10

Site Index

When the site index of a particular species was unknown, the following conversions were used.

Redwood Site = 46.5 + .465(Douglas Fir Site)
Douglas Fir Site = 80.15 + .47(Redwood Site)

Data sources and a statistical summary are described in Research Note No. 5 (Krumland and Wensel 1976). Based on limited data and regression equations described by Wiant (1966), the following relationships were assumed for hardwoods.

Tan Oak Site = 43 + .27(Redwood Site)
= 44 + .20(Douglas Fir Site)

Alder Site = 63 + .28(Redwood Site)
= 63 + .25(Douglas Fir Site)

It should be emphasized that these relationships were developed by graphical comparisons rather than regression methods with actual data.

Breast High Age

Based on the same data set used to develop redwood and douglas fir site index conversions, the following breast high age equations were estimated.

Redwood Age = 8.2+.85 Douglas Fir Age + .046 Douglas Fir Site
Douglas Fir Age = -11.6+.88 Redwood Age + .101 Redwood Site

For hardwoods, Porter and Wiant (1965) estimated that it took 3.2 and 1.2 years for tan oak and alder, respectively, to reach breast high age. Assuming that redwood takes one year and that douglas fir takes seven the following approximations were derived.

Alder Breast High Age = Redwood Breast High Age
= Douglas Fir Breast High Age + 6

Tan Oak Breast High Age = Redwood Breast High Age - 2
= Douglas Fir Breast High Age + 4

Appendix III

Choice of Percentage Points

In the inverse Weibull density function

$$d_i = a + b\{-\ln(1-p_i)\}^{1/c}$$

three separate diameter percentile points (d_1, d_2 and d_3) are needed to solve for the parameters a, b and c. Denote the terms $\{-\ln(1-p_i)\}$ as q_i for convenience. One relationship which is solely a function of the unknown parameter 'c' and can be derived from the three known (predicted) percentile points is

$$(d_1 - d_2)/(d_2 - d_3) = (q_1^{1/c} - q_2^{1/c})/(q_2^{1/c} - q_3^{1/c})$$

The solution for 'c' is indeterminate in this form without further specifying the relationship between the q_i. As the simplest case, we set

$$q_2 = q_1^{x+1}$$

$$q_3 = q_1^{2x+1}$$

where x may take on any value we might assign.

From these restrictions on the q_i we find

$$(d_1 - d_2)/(d_2 - d_3) = q_1^{1/c}(1-q_1^{x/c})/q_1^{((x+1)/c)}(1 - q_1^{x/c})$$

$$= q_1^{-x/c}$$

Hence,

$$c = -x \ln(q_1)/\ln[(d_1 - d_2)/(d_2 - d_3)]$$

$$= -x \ln(-\ln(1 - p_1))/\ln[(d_1 - d_2)/(d_2 - d_3)]$$

In generating the values d_i for estimating the coefficients in the second stage of the model construction, there is a choice in selecting the initial percentage point (p_1) and the value of 'x'. Values of p_1 close to .63 would result in comparable values for p_2 and p_3 regardless of the value for 'x'. Conversely, if p_1 is selected close to 1.00, p_2 is still approximately around .63 but p_3 approaches 0 as 'x' becomes large. A theoretical solution to the optimal choice of p_1 and 'x' was considered but it was abandoned because it was too difficult to deal with. Instead, several possible sets of choices for "x' and p_1 were chosen, the second stage models for d_1, d_2 and d_3 estimated and predicted values of a, b and c were computed. Root mean square differences (RMSD) between predicted values and actual plot parameters were then compared. Values of p_1 = .89 and x = -1.1 appeared to give the smallest RMSD of the possible sets tested although differences between sets were, in general, small.

Appendix IV

Constraints on Predictions

The system of models described previously have been extensively tested under a wide range of species mixes, density and top height (site-age combinations) levels. In some situations which occur largely outside of the ranges of the original data, some of the predictions are illogical. These situations usually occur in young stands at high stocking levels. The principle problem is that the computed values of 'c' or 'a' become negative.

The following correction was found to give fairly satisfactory results.

(1) If the computed value of 'c' was less than .3, it was set at .3.
(2) The value of 'b' and 'a' were computed as usual
(3) If the value of 'a' was less than 0.0, it was set at 0.0 and the value of 'b' was recomputed as:

$$\hat{b} = [(\hat{D})^2 \Gamma(1 + 2/\hat{c})]^{1/2}$$

where

D = predicted quadratic mean DBH under Model I or the actual value under Model II.

$\Gamma(x)$ = gamma function of 'x' (see Appendix I)

With Model II, the parameter 'a' is initially estimated as

$$a = -\hat{b}\Gamma(1 + 1/c) + [\hat{b}^2(\Gamma(1 + 1/\hat{c})^2 - \Gamma(1 + 2/\hat{c})) + D^2]^{1/2}$$

To insure compatibility with the specified quadratic mean diameter, Step (3) above is then followed as a check.