Northern California Forest Yield Cooperative
Department of Forestry and Resource Management
University of California, Berkeley, Ca. 94720

Research Note No. 30 August 28, 1991

The predictive models and procedures used in the Forest Stand Generator (STAG[1])

by

Greg S. Biging, Timothy A. Robards, Paul C. Van Deusen, and Eric C. Turnblom[2,3]

[1] **STAG** is an acronym for Forest **STA**nd **G**enerator for Mixed Conifer Species © U.C. Regents 1986-1991.

[2] Associate Professor, Department of Forestry and Resource Management, University of California, Berkeley, Programmer-Analyst with the California Department of Forestry and Fire Protection, Research Scientist at the USDA Forest Service Southern Forest Experiment Station, New Orleans, Louisiana, and Post Graduate Researcher, Department of Forestry and Resource Management, University of California, Berkeley.

[3] Research conducted under AES projects 3679-ms and 3815-ms. We gratefully acknowledge the important contributions to this work by Peter J. Daugherty, and Vaughn Landrum. We also are grateful to the members of the Northern California Forest Yield Cooperative for their support.

The predictive models and procedures used in the Forest Stand Generator (STAG)

ABSTRACT

The Forest Stand Generator, STAG, is a PC based program that uses statistical routines to produce complete stand descriptions comprised of individual tree measurements of diameter at 4.5 feet above ground (denoted hereafter as diameter at breast height, or DBH), total height, height-to-crown base, species and tree expansion factor. STAG is versatile enough to work with several types of data and still produce complete stand descriptions. This achievement makes it possible to utilize the California Conifer Timber Output Simulator (CACTOS, Wensel and Others, 1986, 1987) for simulation of tree growth and mortality even though the initial datasets could not have been used with CACTOS.

INTRODUCTION

The interior forests of Northern California are typically comprised of mixed conifer species of multiple ages and sizes. Inventory procedures for these lands are varied, as is the experience in the rest of the U.S. There are several common inventory procedures that this work addresses. One typical procedure is to measure DBH and to subsample tree heights and height-to-crown base. This procedure yields what can be considered a "missing data" case. Another common procedure is to record the number of trees by diameter classes. This yields stand table data which is a discrete approximation of the continuous diameter distribution. In some cases only stand summary statistics are recorded such as the basal area per acre (basal area is the cross-section of trees measured at 4.5 feet above ground in square feet on a per acre basis) and number of trees per acre. In these latter two cases no individual tree information is recorded, just overall stand parameters.

A common use of inventory data is for simulation of future growth and yield of stands from which the data were derived. Our goal is to ensure that the three forms of inventory data listed above can be made to conform to the requirements of the California Conifer Timber Output Simulator, CACTOS (Wensel and others 1986, Wensel and Biging, 1987). CACTOS simulates the growth and development of individual trees and requires that species, DBH, tree height (H), height-to-crown base (HCB) or live crown ratio, and tree expansion factor[4] be supplied for each individual tree making up the stand description. When all these data are present we refer to them as a stand description which is comprised of complete individual tree records. To take full advantage of the simulation capacity of CACTOS, these variables should be measured for all trees.

When data sets are not complete STAG can be used to produce complete stand descriptions for a wide class of inventory procedures (Biging, Meerschaert, and Robards 1991). This paper will discuss the estimation procedures used in STAG to (1) fill in missing measurements of tree height, height-to-crown base or both; (2) generate stands from summary statistics; and (3) convert stand table data, numbers of trees by DBH classes and species, to individual tree records so that these stand descriptions (comprised of complete individual tree records) can be analyzed by CACTOS. We also discuss the predictive equations and analytic procedures used to produce complete stand descriptions for these three differing categories of data availability.

---

[4] Tree expansion factor is defined as the number of trees per acre that the sample tree represents.

DATA

Data for this study were provided by the Northern California Forest Yield Cooperative growth and yield project. These data were collected from 710 permanent plots located throughout the mixed conifer region of northern California. Variables measured for each tree included species, DBH, total height, and height-to-crown base. The permanent plots were established in 1978-79 and a five year remeasurement was made in 1983-84. These plots were typically 1/5th acre in size, but contained subplots used to measure submerchantable trees. Usually trees greater than 11.0 inches in DBH were measured on the full plot. Trees between 5.5 and 11.0 inches in DBH were measured on a 1/10th acre subplot and trees between 1.5 and 5.5 inches in DBH were measured on a 1/20th acre subplot. There were some variations in the class limits depending upon the company collecting the data. The five year remeasurement data were used for the models developed in later sections of this paper. Figure 1 shows the location of the permanent plots by township and the Appendix provides summary statistics for much of the data used in this study.

ESTIMATION PROCEDURES

STAG is a PC based program that uses statistical routines to produce stand descriptions comprised of complete individual tree measurements of DBH, total height, height-to-crown base, species and tree expansion factor. There are three main data analysis routines in STAG and distinct statistical procedures used in each corresponding to the three different classes of data availability (filling in missing data, converting stand table data, and generating stands from summary statistics). Each of the three routines are described below. In this section we define estimation techniques for overstory trees. We define overstory trees as those trees greater than a defined threshold value of either 5.5 or 11.0 inches in DBH. The species are denoted throughout the paper using the species codes given in Table 1. These species are classified into 8 different species groups during the simulation process as shown below in Table 2.

Table 1. Species codes and names.

Species Code	Common Name	Species Abbreviation	Scientific Name
01	ponderosa pine	PP	*Pinus ponderosa* (Laws.)
02	sugar pine	SP	*Pinus lambertiana* (Dougl.)
03	incense cedar	IC	*Libocedrus decurrens* (Torr.)
04	Douglas-fir	DF	*Pseudotsuga menziesii* (Mirb.) Franco
05	white fir	WF	*Abies concolor* (Gord. and Glend.) Lindl.
06	red fir	RF	*Abies magnifica* (A. Murr.)
07	lodgepole pine	LP	*Pinus contorta* (Dougl.)
08	white pine	WP	*Pinus monticola* (Dougl.)
09	Jeffrey pine	JP	*Pinus jeffreyi* (Grev. & Balf.)
10	miscellaneous conifers	CM	n.a.
11	chinquapin	CH	*Castanopsis chrysophylla* (Dougl.) A. DC.
12	black oak	BO	*Quercus kelloggii* (Newb.)
13	tan oak	TO	*Lithocarpus densiflorus* (Hook. & Arn.)
14	misc. hardwoods	HM	n.a.

Table 2. Species groups used for modelling in STAG.

Species Group No.	Species Group Name	Sp. Group Abbreviation	Species (and Codes) Included in Group
1	Ponderosa Pine	PP	PP(01), JP(09), LP(07)
2	Sugar Pine	SP	SP(02), WP(08)
3	Incense Cedar	IC	IC(03)
4	Douglas-fir	DF	DF(04), CM(10)
5	White Fir	WF	WF(05)
6	Red Fir	RF	RF(06)
7	Other Hardwoods	OH	CH(11), TO(13), HM(14)*
8	Black Oak	BO	BO(12)

* The OH equations were derived mainly from CH(11), and TO(13)

ESTIMATING TOTAL HEIGHT

STAG can be used to fill in where tree heights, heights-to-crown base, or both are missing provided that the species, DBH, and expansion factors exist for all trees on the plot. Models [1] and [2] are used to estimate missing heights for overstory (> 5.5 inches DBH), and understory trees (≤ 5.5 inches DBH), respectively.

Heights for overstory trees whose diameter exceeds 5.5 inches are estimated as a function of DBH, stand basal area, and elevation as:

[1] $$\hat{H}_O = b_0 + b_1 \cdot \sqrt{DBH} + b_2 \cdot \sqrt{BA_6} + b_3 \cdot E^2$$

where $\hat{H}_O$ = the estimated total height (ft) for overstory trees
BA_6 = the stand basal area (ft^2) in trees greater than 5.5 inches in DBH,
DBH = tree diameter at breast height (DBH > 5.5 inches)
E = stand elevation in feet.

The coefficients b_0, b_1, b_2, and b_3 were estimated for species groups 1-8 (see Table 2) and an all species combined category. Sample sizes for each species ranged from a low of 340 observations for black oak to over four thousand observations on Ponderosa pine and white fir. All standard errors were in the range of 9 - 14 feet. Other model forms which included site index were evaluated, but did not outperform this model. Coefficient values and fit statistics are presented in Table 3.

The model used for predicting heights of understory trees is:

$$[2] \qquad \widehat{H}_u = 4.5 + \frac{\widehat{H}_{5.5} - 4.5}{5.5} \cdot DBH$$

where $\widehat{H}_u$ = the estimated total height (ft) for understory trees whose diameter is in the range $0 < DBH \leq 5.5$

$\widehat{H}_{5.5}$ = the predicted height (ft) of a 5.5 inch DBH tree from equation [1]

Model [2] simply constrains the predicted height of understory trees to be between 4.5 feet and the height of a 5.5 inch DBH tree as predicted by model [1]. We used this constrained equation rather than a statistical model to ensure that the understory height predictions would smoothly join the overstory equation.

Table 3. Coefficients and fit statistics for the total height model [1] for overstory trees.

Species Group and number	$S_{y.x}$	N	b_0	b_1	b_2	b_3
PP [1]	12.144	4173	-38.673	27.073	1.809	-7×10^{-7}
SP [2]	11.215	1070	-36.456	28.328	0.999	-6×10^{-7}
IC [3]	9.406	2260	-28.246	22.713	0.709	-6×10^{-7}
DF [4]	11.488	2458	-34.586	27.400	1.446	-6×10^{-7}
WF [5]	10.700	5167	-40.147	29.353	0.829	-4×10^{-7}
RF [6]	11.397	501	-36.656	28.605	1.005	-5×10^{-7}
OH [7]	13.218	273	-38.731	15.614	2.621	0
BO [8]	14.421	340	-2.386	13.237	1.712	-8×10^{-7}
- All	13.488	16242	-35.36	27.61	1.03	-6×10^{-7}

ESTIMATING HEIGHT-TO-CROWN BASE

To estimate height-to-crown base (HCB) for overstory trees with DBHs > 5.5 inches, a model form based on the logistic equation was chosen so that HCB would be constrained to be between zero and total height. The form of the model selected was:

$$[3] \qquad \widehat{HCB}_O = H \left(1 - e^{-(c_0 + c_1 \cdot \ln BA_6 + c_2 \cdot (DBH/H))^2} \right)$$

where $\widehat{HCB}_O$ = predicted height (ft) to the base of the crown for overstory trees (> 5.5 inches in DBH),
H = is total height (ft),
DBH = diameter at breast height (nearest 0.1 in),
c_0, c_1, c_2 = coefficients estimated for each species group, and
BA_6 is as defined above.

Sample sizes were the same as in estimating the total height model, but standard errors were slightly less ranging between 9 - 11 feet. Coefficients and fit statistics are presented in Table 4.

Table 4. Coefficients and fit statistics for the height-to-crown base model [3] for overstory trees.

Species Group and number	$S_{y.x}$	N	c_0	c_1	c_2
PP [1]	10.375	4173	1.027	-0.112	1.925
SP [2]	9.454	1070	1.222	-0.130	1.400
IC [3]	8.703	2260	1.119	-0.097	0.974
DF [4]	11.140	2458	1.369	-0.162	1.833
WF [5]	10.856	5167	1.298	-0.154	1.831
RF [6]	11.089	501	1.450	-0.160	1.022
OH [7]	9.188	273	1.727	-0.184	0.535
BO [8]	10.315	340	1.313	-0.133	0.745
All	10.580	16242	1.323	-0.146	1.414

Height-to-crown base of understory trees was estimated as:

$$[4] \qquad \widehat{HCB}_u = c_0 + c_1 \cdot DBH + c_2 \cdot H + c_3 \cdot N_6$$

where $\widehat{HCB}_u$ = predicted height (ft) to the base of the crown for understory trees (≤ 5.5 inches in DBH),
H = is total height (ft),
N_6 = number of trees per acre with DBH > 5.5 inches, and
c_0, c_1, c_2 = coefficients estimated for each species group.

Table 5. Coefficients and fit statistics for the height-to-crown base model [4] for understory trees.

Species Group and no.	$S_{y.x}$	N	c_0	c_1	c_2	c_3	b
PP [1]	3.393	1377	2.727	1.737	0.166	-0.0181	1.3971
SP [2]	3.512	224	4.214	1.110	0.252	-0.0192	1.2756
IC [3]	3.097	996	1.764	0.894	0.197	-0.0069	1.5468
DF [4]	4.753	960	1.659	2.567	0.188	-0.0168	2.0209
WF [5]	3.901	2470	0.866	1.468	0.295	-0.0129	1.9335
RF [6]	3.741	131	3.361	0.437	0.294	-0.0132	2.0591
OH [7]	3.909	75	9.145	-0.599	0.411	-0.0165	1.0384
BO [8]	4.210	90	1.290	-0.032	0.447	-0.0153	2.2357
All	3.988	6323	1.922	1.201	0.302	-0.0159	1.9789

For model [4] we observed that variance increased with increasing predictions of height-to-crown base. We formulated a simple model for this relationship as:

[4b] $$\sigma_{iu}^2 = b \cdot \widehat{HCB}_{iu}$$

where $\widehat{HCB}_{iu}$ = predicted height (ft) to the base of the crown for the i^{th} understory tree ($\leq$ 5.5 inches in DBH),

σ_{iu}^2 = the variance around the regression of the height-to-crown base model [4] for the i^{th} understory tree (i = 1 to n)

b = a coefficient estimated for each species group

The procedure for adding stochastic errors is discussed in more detail in a later section. Briefly, we predict height-to-crown base for understory trees with DBHs $\leq$ 5.5 inches using equation [4]. Stochastic errors are then added to the prediction. We assume the stochastic errors are distributed normally with mean zero, and variance as given in equation [4b]. The estimated values of beta (b) in equation [4b] are given in the Table 5 above.

With these equations it is possible to "fill in" or estimate missing values of height and height-to-crown base for individual trees. The only exogenous variable that needs to be supplied for each stand is elevation. Basal area (BA_6) can easily be computed directly by summing the per acre individual tree basal areas obtained from the individual tree DBHs and expansion factors contained in the stand description file. Number of trees (N_6) can easily be calculated from the expansion factors associated with individual trees in the stand description file.

CONVERTING MERCHANTABLE HEIGHTS TO TOTAL HEIGHTS

The four different types of tree height measurements allowed in STAG include: 1) total heights, 2) heights to a merchantable top (≤6.5 in. d.i.b.), 3) heights measured to whole (16.5 ft.) logs, or 4) heights measured to half logs (8.25 ft.). Within a STAG stand description file comprised of individual tree measurements all heights must be of the same measurement standard. CACTOS requires total heights for individual trees, but STAG can manipulate merchantable height to obtain an estimate of total height. STAG uses a taper equation to solve for total height for the six major conifer species (species group numbers 1-6 in Table 2) whenever height to a merchantable top or number of 16.5 ft. logs is supplied[5] .

The equation used to convert merchantable height (MH) to total height (H) is derived from a sigmoid taper equation (Biging, 1984). The total height estimate obtained from inverting the taper equation is:

[5] $$\widehat{H}' = \frac{MH \cdot (\lambda)^3}{(1 - \exp[(d/DBH - b_1) / b_2]]^3}$$

where $\widehat{H}'$ = the predicted total tree height (ft.) estimated from merchantable height
λ = 1 - $\exp(-b_1/b_2)$
DBH = the diameter at breast height (in.)
d = the merchantable top diameter (≤ 6.5")
MH = the height to the merchantable top diameter (ft.)
exp(x) = 2.71828... raised to a power of 'x', and
b_1,b_2 are species specific coefficients given in Table 6

Table 6. Coefficient estimates by species for equation [5] from Biging (1984).

Species and spp. codes	N	b_1	b_2
PP [1]	2014	1.019589	0.335666
SP [2]	692	1.06932	0.415632
IC [3]	541	1.071343	0.472157
DF [4]	1588	1.029288	0.334012
WF [5]	2645	1.092615	0.365295
RF [6]	312	1.075880	0.353784

[5] The height conversion process is not intended to encourage the measurement of other than total heights. Rather, it is intended to allow the use of older inventory data.

If the heights of trees are entered as number of logs, the program first converts these to heights to the given merchantable top using equation [6], and then uses equation [5] to predict total heights. The equation to estimate height to the merchantable top (MH) when only the number of logs is known is given by:

$$[6] \qquad \widehat{MH} = \left[SH + \frac{LL}{2}\right] + NLOGS \cdot LL$$

where
$\widehat{MH}$ = the estimated height to the merchantable top diameter (ft.)
SH = stump height of the tree = 1.5 ft.
LL = log length in feet (16.5 or 8.25 feet)
NLOGS = the number of logs of length LL for a tree

STOCHASTIC ERRORS

When filling in missing data (height or height-to-crown base) or generating stands from summary statistics (discussed in a later section) the user can either make a deterministic or a stochastic prediction of missing values. Choosing stochastic errors means that a random value will be added to the prediction to reflect that an individual tree's dimensions cannot be predicted with certainty. Thus, a random value will be added or subtracted from the prediction. The random value is drawn from a normal distribution with mean zero and variance equal to the estimated variance around the regression ($S^2_{y \cdot x}$) as given in Tables 3 and 4. In the case of the understory height-to-crown base model [4], the distributional mean is zero, but the variance is proportional to the predicted height-to-crown base (see equation [4b]). If random errors are not requested, then the missing value is set equal to the model prediction (deterministic prediction). If random errors are not added, all predicted heights and crown ratios will be identical for a given diameter of a particular species, given that basal area and elevation are the same.

PARAMETER UPDATING

If the user wants to incorporate knowledge of a local sample into the height model coefficients a Bayesian update of the first two parameters of the height model [1] is possible. Alternatively, an *ad hoc* weighting scheme patterned after the linear composite estimators (Burk and others 1982) can be chosen. In both cases only the first two parameters (b_0, and b_1) are allowed to be updated because the effects of elevation (E) and density (BA_6) can not be adequately described with a local sample.

The *ad hoc* approach adjusts the amount of change to the model parameters by a constant ratio (k) between 0 and 1. A weight of k equal to zero causes the update routine to abort (no update), while a weight of k equal to one places all the emphasis on the local sample to determine the coefficient values to be used for the height prediction equations. This *ad hoc* weighting process is given as:

[7] $$\hat{\beta} = K \cdot I \cdot \beta_L + (I - K) \cdot \beta_D$$

where $\hat{\beta}$ = { b_0^u, b_1^u } where b_0^u and b_1^u are the updated parameter estimates

K = the *ad hoc* weight ($0 \le k \le 1$)
I = an identity matrix (2x2)
β_D = the database estimate of the parameters (2x1)
β_L = the estimate of the parameters based upon the local sample (2x1)

We modified the true Bayesian method because in prior work (Van Deusen, 1984) we found that it worked poorly. With relatively small sample sizes the Bayesian update could result in large covariance terms in the local covariance matrix which could cause the updated parameter estimates to behave poorly. For example, the local parameter estimates could indicate that both the data base slope and intercept coefficients should be increased over their database couterparts. A large negative covariance term in the local covariance matrix could force these two coefficients to move in opposite directions, regardless of the fact that both local parameter estimates were larger than the database estimates. Because of this we modified the Bayesian approach and have termed it a pseudo-Bayesian approach. The main difference between a Bayesian and a pseudo-Bayesian approach is that for the latter we utilize only the variance terms in the variance-covariance matrices of the local and database samples to avoid problems associated with the covariance terms.

The pseudo-Bayesian approach is more conservative than the ad-hoc procedure. If the local sample is small then the updated coefficients for the height prediction equation are quite close to the database values. If, however, there is a large local sample, then the pseudo-Bayesian estimates are a compromise between the database values and those determined from the local sample. The pseudo-Bayesian update is given by:

[8] $$\hat{\beta} = W \cdot \beta_L + (I - W) \cdot \beta_D$$

where $\hat{\beta}$ = { b_0^u, b_1^u } where b_0^u and b_1^u are the updated parameter estimates

I = an identity matrix (2 x 2)
β_D = the database estimate of the parameters (2x1)

β_L = the estimate of the parameters based upon the local sample (2x1)

W = the weighting matrix (2x2) = $(V_D^{-1}+V_L^{-1})^{-1}\ V_L^{-1}$

V_D^{-1} = the diagonal elements of the inverse of the variance matrix for the database parameters

V_L^{-1} = the diagonal elements of the inverse of the variance matrix for the parameters based upon the local sample

Van Deusen (1984) found that if the local estimate is of sufficient size it is often the best, but when uncertainty exists the *ad hoc* or pseudo-Bayes methods are reliable, with the pseudo-Bayes being conservative and of low risk.

GENERATING STANDS FROM SUMMARY STATISTICS

In cases where no individual tree measurements are available or when only summary statistics are recorded by species it is possible to generate a facsimile stand description comprised of complete individual tree records based upon the summary statistics. With the knowledge of the summary statistics it is possible to generate a diameter distribution as developed in a later section. Individual tree diameters can be sampled from this distribution. Tree height and height-to-crown base values are estimated from equations [1] through [4] to complete the facsimile stand description. The goal of this methodology is to produce a facsimile stand description of complete individual tree records that is plausible given the specified summary statistics.

Generation of the Overstory of Trees

We will define overstory trees as those greater than a defined threshold value (usually 5.5 or 11.0 inches in DBH), and understory trees as those at or below the threshold value. We have developed separate approaches for generating overstory and understory trees to achieve better accuracy in predicting missing data values.

The joint distribution of species, DBH, height, and height-to-crown base is formulated as a product of probability density functions (Van Deusen 1984, and Biging and Wensel 1987). This joint probability distribution for these overstory trees can be represented as a mixture of distributions as:

[9] $$p(D,H,HCB) = \sum_{s=1}^{S} p(\text{Species})\cdot p(\text{DBH} \mid \text{Species})\cdot p(H \mid \text{Species, DBH})\cdot p(\text{HCB} \mid \text{Species, DBH, H})$$

where S = the number of species present in the stand

The joint probability distribution of DBH, total height and height-to-crown base [p(D,H,HCB)] is factored as a product of three conditional distributions. The first term on the right hand side is p(Species) which is the fraction of each species in the stand. This is easily specified by the user by supplying the number of trees per acre by species in the hypothetical stand.

The three conditional distributions are for DBH, height and height-to-crown base. The first of these conditional distributions is that of diameter given species [p(DBH|species}. The conditional diameter distribution can be generated from either a two-parameter truncated Weibull or a negative exponential distribution by relating the summary statistics to the parameters of these distributions. The first two moments of the Weibull distribution correspond to the average diameter of the species, and the quadratic mean diameter^2 of the species which can be derived from basal area and number of trees for each species (see equation [13]).

We found that the first moment (average diameter for a given species) could be accurately predicted as a function of the quadratic mean stand diameter, elevation and numbers of trees in the species. This is discussed more fully in a following section (see equation [12]). The user can generate a diameter distribution for each species having knowledge of only the number of trees and basal area in each species. Individual tree DBHs are then randomly generated using an inverse transformation method for either the two-parameter Weibull or the negative exponential.

To randomly sample from this distribution we will consider it a probability density function, compute it's associated cumulative distribution function, and finally compute the inverse cumulative distribution function from which we may generate DBHs. Since the cumulative distribution function produces a probability and by definition probabilities are bounded between 0 and 1 we may use uniformly distributed random deviates bounded between 0 and 1 to generate values for input into the inverse cumulative distribution function.

A second distribution, the negative exponential, was provided for the infrequent case where a balanced uneven aged condition exists within a stand. The details for the procedure of fitting the distribution and a list of the necessary stand summary statistics are provided in a section below.

With either the Weibull or negative exponential distribution only unimodal distributions can be generated for a given species. In most cases there are too few trees of a given species to develop more complex distributional models. However, because we allow each species to have its own diameter distribution it is possible to build multi-modal distributions for a stand.

Once the diameters are specified with the diameter distribution the height and height-to-crown base values are predicted with equations [1] and [3]. These later two equations correspond to the conditional probability distributions of p(H | Species, DBH) and p(HCB | Species, DBH, H), respectively. Elevation also needs to be supplied since it is an independent variable in the height prediction equation [1]. Random stochastic errors distributed as $N(0,S^2_{y \cdot x})$ for equation [1] or $N(0, b\,\hat{y}_i)$ for equation [3] are added to the predictions if the random error feature has been selected.

There are alternatives to the factorization approach utilized in this study. For example, the joint distribution of diameter, height, and height-to-crown base could have been modeled as a trivariate distribution. We did not investigate this latter approach because we had relatively few measured trees on each of the remeasured permanent plots (usually less than 20). With a factorization approach there is the additional advantage that any number of species can be modelled.

Weibull Distribution

The Weibull distribution has been widely used in forestry applications for describing the diameter distributions of stands. This use stems both from the Weibull's shape and ease of estimation of parameters. Since the purpose of generating hypothetical individual tree data from stand summary information is for use in CACTOS to project future yields, the data should be compatible to allow for transfer from STAG to CACTOS. Thus a truncation of 5.5 inches DBH is used since the CACTOS growth models are fit on data greater than 5.5 inches DBH. The three parameter Weibull may be reduced to the two parameter Weibull since the location parameter, typically called *a*, is zero. The two parameter truncated Weibull density function is given as (Van Deusen 1984):

[10] $$f(x) = \left[\frac{c}{b}\right] \cdot \left[\frac{x}{b}\right]^{c-1} \cdot e^{\left[(T^c - x^c) \cdot b^{-c}\right]}$$

where $f(x)$ = frequency of trees in diameter class x,
x = midpoint DBH of diameter class; $x \geq T$,
T = truncation DBH (5.5"),
b and c = parameters > 0.

<u>Deriving the *b* and *c* coefficients</u>

To specify a particular distribution from this Weibull family, we need to define the b and c parameters. The moment equation for the two parameter truncated Weibull is given as:

$$[11] \qquad Ex^r = b^r \cdot e^{(T/b)^c} \cdot \left[\Gamma(r/c+1) - c \cdot b^{-(r+c)} \cdot \int_0^T x^{r+c-1} \cdot e^{-(x/b)^c} \cdot dx \right]$$

where Ex^r = the expectation of the r^{th} moment of x or DBH,
$\Gamma(r/c + 1)$ = the gamma function of (r/c + 1).

The first and second moments are used to simultaneously solve for the *b* and *c* parameters. It is known from Cauchy's inequality that the arithmetic mean must be less than or equal to the quadratic mean stand diameter which is the square root of the equation [13] below. The arithmetic mean stand diameter is predicted as a fraction of the quadratic mean stand diameter where the fraction is constrained to be less than 1 through use of the logistic function. The arithmetic mean stand diameter is predicted as follows:

$$[12] \qquad D^{(1)} = \overline{DBH} = \left[\beta_0 + \frac{(1-\beta_0)}{\left(1+e^{\left(-\beta_1 - \beta_2 E - \beta_3 \ln(\overline{D}_q) - \beta_4 N_6^{-1} - \beta_5 \overline{D}_q^{-1}\right)}\right)} \right] \cdot \overline{D}_q$$

where $D^{(1)}$ = the estimated $\overline{DBH}$ = the first moment or mean stand diameter for a given species

$\overline{D}_q$ = quadratic mean diameter of trees for a given species > 5.5" DBH,
N_6 = number of trees for a given species > 5.5" DBH,
E = elevation (ft.), and
$b_0, \ldots, b_5$ = the coefficients estimated from regression (see Table 7).

Table 7. Coefficients and fit statistics for the mean stand diameter model for species number 1-8 combined.

Species number	N	MSE	β_0	β_1	β_2	β_3	β_4	β_5
[1-8]	2078	0.363	0.75637	-12.12687	-0.00018	3.62041	6.15495	56.31421

The second moment is the quadratic mean stand diameter squared which is given by definition as:

$$[13] \qquad D^{(2)} = \overline{D}_q^2 = \frac{1}{n} \sum DBH_i^2 = \frac{SBA_6}{K' \cdot SN_6}$$

where $D^{(2)}$ = the estimated second moment or quadratic mean stand diameter squared ($\bar{D}_q^2$) of a given species,
DBH_i^2 = DBH^2 of the i^{th} tree
n = the number of trees on a plot
K' = 0.005454 which is a conversion factor for basal area in square feet to diameter squared in square inches.
SBA_6 = basal area of trees for a given species > 5.5" DBH,
SN_6 = number of trees for a given species > 5.5" DBH

The absolute difference between the two moments given above and the predicted moments given an estimated *b* and *c* are combined into an overall error figure. This figure must be less than 5% (E%) of the respective moments. The error formula is:

[14] $$E = \frac{E\% \cdot D^{(1)} + \sqrt{E\% \cdot D^{(2)}}}{2}$$

where E% = the percent error allowed in estimating the moment, and
$D^{(1)}, D^{(2)}$ are defined as above.

Using ponderosa pine as an example, if the species basal area were given as 150 ft^2 and the species number of trees per acre given as 300 then the maximum allowable error in finding the Weibull parameters would be:

$$E = \frac{0.05 \cdot 9.82 + \sqrt{0.05 \cdot 91.68}}{2} = 1.32 \text{ inches,}$$

where 9.82 is $D^{(1)}$ and 91.68 is $D^{(2)}$. When estimates of *b* and *c* are obtained then predicted values of $D^{(1)}$ and $D^{(2)}$ are obtained and summed as:

$$\widehat{E} = \frac{|9.82 - \widehat{D^{(1)}}| + \sqrt{|91.68 - \widehat{D^{(2)}}|}}{2}$$

where $\widehat{D^{(1)}}$ and $\widehat{D^{(2)}}$ are the predicted first and second moment respectively. If $\widehat{E}$ is less than E then the estimates of *b* and *c* are close enough and the procedure stops, otherwise new estimates of the parameters are calculated and the process is repeated.

To begin the algorithm for determining the coefficients two predictive equations are used to provide starting values for the parameters. These equations were fit using multiple linear regression for converged values.

$$\hat{b} = -1.260 + 1.183 \cdot D^{(1)} - 0.0018 \cdot D^{(2)}$$

$$\hat{c} = 8.112 - 0.543 \cdot D^{(1)} + 0.013 \cdot D^{(2)}$$

Next a Newton-Raphson procedure is used to minimize $\hat{E}$. If the Newton-Raphson fails to converge within 6 iterations then a grid search is used to minimize $\hat{E}$. The Newton-Raphson procedure is much faster than a grid search and is quick to converge for most stands with mean stand diameters of about 14" DBH and above. The Newton-Raphson method generally does not provide stable estimates of the change in the *c* parameter necessary to minimize $\hat{E}$, and thus influences almost exclusively the *b* parameter. Thus if the starting value for *c* is not close to optimum then a grid search will be necessary.

Multiple minimums exist for the error surface, $\hat{E}$, with respect to *c* while only one minimum exists with respect to *b*. Thus a decision to restrict the *c* parameter to its lowest minimum was made in order to produce the maximum variance for the distribution. This was done after examining permanent plot data for Northern California and noting the large variability over diameter classes which exists on a plot to plot basis. If in the future these stands approach a more even-aged structure then the program can be adjusted to reflect this by restricting the *c* parameter into the next largest minimum, thus reducing the variability over size classes.

The grid search algorithm begins by searching for the minimum error over a course grid (increment of 0.5) with respect to $\hat{b}$ and $\hat{c}$ This course grid search is accelerated by retrieving the $\widehat{D^{(1)}}$ and $\widehat{D^{(2)}}$ from two binary files. Next the range of the parameters are reduced to be around the minimum found in the course grid search, the increments for $\hat{b}$ and $\hat{c}$ are reduced to a third of their previous value, and a finer search is performed. Up to ten iterations of increasingly finer grid searches are performed. As with the Newton-Raphson technique the convergence criteria is that $\hat{E}$ be less than E. Once the grid search converges a fine tuning is performed where $\hat{b}$ and $\hat{c}$ are adjusted slightly so that the error in estimating the first moment is approximately the same as that of the second moment.

Estimating DBHs

The derivation for the inverse cumulative distribution function of the truncated two-parameter Weibull is as follows. The probability density function is integrated from the lower truncation point (T) to the diameter of interest (x).

$$[15] \qquad F_T(x) = \int_T^x \left[\frac{c}{b}\right] \cdot \left[\frac{x}{b}\right]^{c-1} \cdot e^{\left[(T^c - x^c)\cdot b^{-c}\right]} dt \quad = \left[1 - e^{(T/b)^c} \cdot e^{-(x/b)^c}\right]$$

where $F_T(x)$ = the cumulative number of trees between the lower truncation point (T), and the specified upper diameter (x)

$\hat{b}$ and $\hat{c}$ = estimated constants

This provides the cumulative number of trees up to the the diameter of interest. Generating a uniform random number between 0 and 1 and multiplying it by the total number of trees gives us a value for F, we may then solve for the DBH by inverting the above equation.

$$[16] \qquad DBH = b \cdot \left[(T/b)^c - \log_e (1 - F_T(x))\right]^{1/c}$$

Thus by generating a uniform random number ($F_T(x)$) we can use the inverse transform of the cumulative distribution function to estimate diameters at breast height.

Negative Exponential Distribution

The diameter distributions of "balanced" uneven aged stands (Meyer 1952) are often characterized as being distributed according to the negative exponential distribution. A typical method for applying the distribution to a stand is with the diminution quotient or "Q" value (Husch, Miller and Beers 1982, Davis and Johnson 1987). To obtain the number of trees in the next to the largest diameter class we would simply multiply Q by the number of trees in the largest diameter class. Thus for the next smallest diameter class we would multiply Q times the number of trees in the next largest diameter class or Q^2 times the number of trees in the largest diameter class. To compute the number of trees in each diameter class we need to specify Q, a range of tree diameters, their diameter class (i.e. 2"), and the number of trees in the largest diameter class. Unlike with the truncated Weibull distribution, we use the negative exponential distribution to simultaneously generate both overstory, and understory trees.

The negative exponential distribution is given as follows:

$$[17] \qquad \widehat{SN}_n = k \cdot e^{-a\, DC_n}$$

where $\widehat{SN}_n$ = the estimated number of trees of a given species in n^{th} diameter class DC_n,
DC_n = the n^{th} diameter class, and
a and k = coefficients.

To specify the distribution we need to define k and a which can be done by using Q, a value which may be more meaningful to a manager than the a and k coefficients of the negative exponential function.

Deriving the a and k coefficients

The coefficient a is derived from the definition of Q:

[18] $$\widehat{Q} = \frac{\widehat{SN}_{n-1}}{\widehat{SN}_n} = \frac{k \cdot e^{-a\,DC_{n-1}}}{k \cdot e^{-a\,DC_n}} = e^{a \cdot C}$$

where $\widehat{Q}$ = the estimated diminution quotient
SN_{n-1} = the number of trees for a species in the next to the largest diameter class,
SN_n = the number of trees for a species in the largest diameter class,
DC_{n-1} = the next to the largest diameter class of a given species,
DC_n = the largest diameter class of a given species,and
C = the size in inches (width) of the diameter class of a given species.

Solving for a we get

[19] $$\hat{a} = \frac{\log \widehat{Q}}{C}$$

Since we know the number of trees in the largest diameter class and the a parameter we may solve for k using the negative exponential equation,

[20] $$\widehat{k} = \frac{\widehat{SN}_n}{e^{-a\,DC_n}}$$

So we can see that when Q and SN_n are known we can estimate the a and k parameters needed for the negative exponential distribution in equation [17].

Calculations when Q or SN_n are unknown

If either Q or SN_n are unknown then the basal area for the species on the plot (SBA) is used to compute the missing variable.

If Q and species basal area (SBA) are known, but SN_n is unknown we iteratively solve for SN_n using equation [21]. In equation [21] species basal area is formulated as the sum of the number of trees in a diameter class multiplied by the square of the diameter class.

$$[21] \qquad SBA = K' \cdot \sum \widehat{SN}_i \, DBH_i^2$$

where
SBA = total basal area for the species on the plot in square feet,
$\widehat{SN}_i$ = the estimated number of trees of the species in the i^{th} diameter class,
DBH_i = the midpoint DBH of the i^{th} diameter class in inches for the species,
K' = a constant (0.005454) for converting basal area from square inches to square feet.

To estimate SN_n from Q and SBA we initially give SN_n a starting value of one. Then if the estimated SBA is less than the specified SBA, SN_n is increased by 0.001 or visa versa. Using the new estimate of SN_n the procedure is repeated until the difference between the specified and estimated SBA is less than one square foot. Q can then be estimated from equation [18].

If Q is unknown then SBA and SN_n must be given. Q is then computed using an iterative process where Q is initially set to 1.1. The number of trees in each diameter class *i* of a given species is estimated by

$$[22] \qquad SN_i = SN_n \cdot Q^{n-i}$$

SBA is then estimated and compared to the specified basal area and the estimate of Q is incremented identically as the estimated SN_n is incremented above. The same threshold of one square foot of basal area difference between the specified and estimated SBA is used as a stopping criteria.

Now that all of the necessary information is complete the coefficients for the negative exponential distribution may be easily computed. The diameters are simulated and written to the stand description file with an expansion factor of one. The tree total height and height-to-crown base is also estimated given the simulated diameter at breast height using equations [1] to [4]. The total number of trees for the species on the plot will be rounded to the nearest integer so that all the trees in the completed stand description file will have an expansion factor of one.

Estimating DBHs

The derivation for the inverse cumulative distribution function of the negative exponential is as follows. The probability density function is integrated over the range of diameters up to the diameter of interest.

$$[23] \qquad F(DBH) = \int_{m}^{DBH} \frac{\hat{k}}{C} e^{-\hat{a} \cdot DC} \cdot dDC = \frac{\hat{k}}{C} \left[\frac{e^{-\hat{a}\,DC}}{-\hat{a}}\right]_{m}^{DBH} = -\left[\frac{\hat{k}}{C \cdot \hat{a}}\right] \cdot \left[e^{-\hat{a} \cdot DBH} - e^{-\hat{a} \cdot m}\right]$$

where $F(DBH)$ = the cumulative number of trees between the minimum diameter (m), and the specified upper diameter of interest (DBH),

$\hat{k}$ and $\hat{a}$ = estimated constants,

m = $DC_{min} - \frac{C}{2}$, with C = the class width in inches,

DC_{min} = minimum diameter class (in).

This provides the cumulative number of trees up to the the diameter of interest. Generating a uniform random number between 0 and 1 and multiplying it by the total number of trees gives us a value for F(DBH), we may then solve for the DBH by inverting the above equation.

$$[24] \qquad DBH = \frac{-\log_e\left(-F(DBH) \cdot \left(\frac{C \cdot \hat{a}}{\hat{k}}\right) + e^{-\hat{a} \cdot m}\right)}{\hat{a}}$$

Thus by generating a uniform random number (F) we can use the inverse transform of the cumulative distribution function to estimate diameters at breast height.

Generation of the Understory Trees

As an adjunct to the stand generation techniques (overstory generation) we have developed the capability to generate understory trees. The understory trees can be between 1.0 and 11.0 inches at DBH. The overstory of trees (measured or generated) can be used to predict the b and c parameters of the Weibull needed to generate understory trees. Understory tree height and height-to-crown base values are estimated from equations [2] and [4] to complete the understory facsimile stand description. Because stands of trees are often simulated for over 30 years with the CACTOS system it is essential to be able to generate an understory component that matures with relatively long simulations. One reason we separated the overstory and understory components is that there is much greater variability (plot to plot, or stand to stand) in the number of understory trees than in the number of overstory trees. Hence the understory generator is inherently more imprecise.

A two-parameter Weibull distribution was fit to the understory component (1.0" ≤ DBH ≤ 11.0") for each of the 308 permanent plots for which there were at least 6 understory trees present on the plot with a plot average diameter exceeding 5.5 inches. Six trees was chosen as a minum number needed for estimating the 2 parameters of the Weibull distribution, although most plots had many more than 6 trees. An 11.0" DBH was chosen as an upper value for the distribution rather than a 5.5" DBH value to allow for a more regular distributional form and to increase the number of trees available for modelling the understory diameter distribution. Even though an upper DBH value of 11.0 inches was chosen, the understory generation can be specified for any range within these limits. Summary statistics for the understory component of the 308 permanent plots used to model the Weibull diameter distribution are presented in Appendix B.

We found that the coefficients of the Weibull for the understory could be predicted as the following functions of overstory parameters:

[25] $$\hat{b} = b_0 + \frac{b_1}{N_6} + \frac{b_2}{DBH_{min}} + b_3 CV_6 + \frac{b_4 CV_6}{DBH_{min}}$$

[26] $$\hat{c} = c_1 \exp(c_2\hat{b} + c_3 X^{c_4} + c_5 Y + c_6 Z)$$

where

$\hat{b}$	= predicted value for b (scale) parameter of two parameter left truncated Weibull
$\hat{c}$	= predicted value for c (shape) parameter of two parameter left truncated Weibull
N_6	= the number of trees per acre greater than 5.5 inches in DBH,
DBH_{min}	= the minimum diameter measured on a specific plot, usually 1.0 or 2.0 inches
CV_6	= the coefficient of variation of DBH for trees greater then 5.5 inches DBH
X	= $0.75 + BA_6/87.945 - SDI_6/131.0$
BA_6	= Stand basal area in trees greater than 5.5 inches DBH
SDI_6	= Stand density index considering only trees greater than 5.5 inches DBH (cf. Reineke 1933, Avery and Burkhart 1983)
Y	= $0.035 + 1/(\hat{b} - SDI_6)$
Z	= $410.0 + 1/[\ln(\hat{b}) - \ln(BA_6)]$
$b_0, \ldots, b_4$	= are b-coefficients estimated for all species combined
$c_1, \ldots, c_6$	= are c-coefficients estimated for all species combined

Coefficient values and fit statistics appear in Table 8 below. Due to the very great inherent variability of the understory component, these predictive equations explain a small but significant portion of the total variability. Therefore the predicted parameters resulting from using these equations will not be very precise but are still preferred over using a simple average value. In general, predicting the parameters of a Weibull distribution from stand characteristics even in situations where there is not a great deal of variation has proven difficult, and R^2 values are typically less than 0.10 (Knoebel and Burkhart 1991).

Table 8. Coefficients and fit statistic for the understory Weibull parameters for all species combined estimated on n = 308 plots.

MSE	b_0	b_1	b_2	b_3	b_4	
3.718	15.4269	-66.160	-15.0426	-0.13859	0.22027	
MSE	c_1	c_2	c_3	c_4	c_5	c_6
3.691	0.57718	0.31116	0.51325	2.9000	-18.8181	4.0501×10^{-5}

Specification of total numbers of understory trees

Even though the *b*, and *c*-parameters of the Weibull distribution can be directly predicted via the two equations listed above, this distribution will give only the relative frequency of tree sizes. Therefore, the total number of understory trees needs to be specified before the understory can be generated. The total number of understory trees can also be predicted from overstory parameters.

Predicting the total number of understory trees from overstory parameters is analogous to predicting the number of ingrowth trees (numbers of trees that will reach some minimum surveyed size in a specified time) with some obvious differences. They are similar in that both involve only use of overstory conditions to estimate the condition of the understory. Prediction of ingrowth numbers is arguably more well defined than prediction of total understory numbers in the sense that one particular size class is under scrutiny while prediction of understory numbers may involve a broad spectrum of size classes. On the other hand, estimates of ingrowth are further complicated by an implied growth rate of trees whose exact sizes are unknown, while estimates of the total number of understory trees represent a static depiction of the stand at one instant in time. Both estimation problems are complicated by the fact that stands currently with similar overstory conditions may have had dissimilar histories, which may result in dissimilar understory conditions.

Models frequently used to predict ingrowth have been reviewed by Shifley (1990). Typically, variables important to the prediction of ingrowth involve stand density measures such as basal area per acre, number of trees per acre, percent stocking, and sum of diameters per acre. These variables also affect growth rates of individual trees, so their superiority in predicting ingrowth is somewhat to be expected. One might also expect that additional variables may be required to predict total number of understory trees due to the previously noted differences between these two estimation problems.

A useful approach to modelling the number of understory trees was found by viewing the problem as specification of total stand structure based upon what was found in the overstory portion only. This lead to the investigation of several stand structure variables. Shifley and Lentz (1985) pointed out that the ratio of the mean DBH to the standard deviation of DBH was a valuable index to the *c*, or shape, parameter in the Weibull distribution. Miller and Weiner (1989) and Knox and others (1989) found that the inverse of Shifley's index, commonly known as the coefficient of variation, was useful in describing "size inequality", or the degree of size hierarchy development in populations of forest trees. We found that the ratio of variance of DBH to the mean DBH was a useful predictor in our models for estimating total number of understory trees.

A model for predicting the number of understory trees was patterned after the ingrowth models of Ek (1974) and Hyink and Moser (1983). The same model form is used for predicting the number of trees between 1.5 and 5.5 inches DBH ($N_{1\text{-}6}$), as for the number of trees between 5.6 inches and 10.5 inches DBH ($N_{6\text{-}11}$). The predictions for understory tree numbers are given by:

[27] $$N_{1\text{-}6} = \exp\{\, b_0 + b_1 \cdot DSUM_6^{b_2} \cdot N_6^{-1} + b_3 \cdot (R_6 + 1.5) \cdot N_6^{b_4} \,\}$$

[28] $$N_{1\text{-}11} = \exp\{\, c_0 + c_1 \cdot DSUM_{11}^{c_2} \cdot N_{11}^{-1} + c_3 \cdot (R_{11} + 1.5) \cdot N_{11}^{c_4} \,\}$$

where
- $\widehat{N}_{1\text{-}6}$ = the predicted number of trees per acre with $1.5 \leq DBH \leq 5.5$
- $\widehat{N}_{1\text{-}11}$ = the predicted number of trees per acre with $5.5 < DBH \leq 10.5$
- N_6 = the number of trees per acre whose DBH > 5.5 inches
- N_{11} = the number of trees per acre whose DBH > 10.5 inches
- R_6 = the ratio of variance of DBH to mean DBH for trees > 5.5 inches
- R_{11} = the ratio of variance of DBH to mean DBH for trees > 10.5 inches
- $DSUM_6$ = the sum of the diameters for trees whose DBH > 5.5 inches
- $DSUM_{11}$ = the sum of the diameters for trees whose DBH > 10.5 inches
- $b_0, \ldots, b_4$ = b-coefficients estimated for each forest type (see Table 9a), and
- $c_0, \ldots, c_4$ = c-coefficients estimated for each forest type (see Table 9b)

We found that predictions could be improved through stratification by forest type. An analysis was performed to see if any of the classes could be combined, but we found that a statistically significant improvement was made by using separate coefficients for each major forest type for predicting both

$N_{1\text{-}6}$ and $N_{6\text{-}11}$. The coefficients were therefore estimated by timber type and are given in Tables 9a and 9b.

Table 9a. Coefficients for $N_{1\text{-}6}$

Timber type	$S_{y.x}$	N	b_0	b_1	b_2	b_3	b_4
Douglas-fir	52.8	25	9.015	-0.412	1.0	0.00091	1.0
Mixed conifer	128.0	469	6.579	-0.211	1.0	0.00110	1.0
Ponderosa pine	135.0	59	6.018	-0.002	1.691	0.00306	1.0
True fir	81.8	83	7.100	-0.266	1.0	0.00114	1.0

Table 9b. Coefficients for $N_{6\text{-}11}$

Timber type	$S_{y.x}$	N	c_0	c_1	c_2	c_3	c_4
Douglas-fir	35.9	25	6.525	-0.189	0.993	0.00223	1.0
Mixed conifer	50.0	468	6.501	-0.146	1.0	0.00245	0.870
Ponderosa pine	63.4	56	6.675	-0.188	1.0	0.00324	1.0
True fir	37.5	83	7.252	-0.626	0.836	0.00008	1.0

STAG automatically determines which forest type the stand description belongs to using the following rules:

Timber type	Definition
Douglas-fir	Douglas-fir comprises ≥ 80% of the stand basal area (BA_6 or BA_{11})
Ponderosa pine	ponderosa pine comprises ≥ 80% of the stand basal area (BA_6 or BA_{11})
True fir	red fir and white fir comprise ≥ 80% of the stand basal area (BA_6 or BA_{11})
Mixed conifer	no one species (PP, SP, DF, WF, RF, IC) exceeds 80% of the stand basal area (BA_6 or BA_{11})

It should be noted that these models are accurate, but not precise. That is to say, there is a large variance associated with these predictions. Therefore, the user is given two options for specifying the number of understory trees. The first option is prediction of understory tree number using equations [27] and [28]. This predicted number of understory trees for a given stand specification is displayed so that the user can accept the model prediction, or specify another value in lieu of the predicted number. This second option is provided for cases in which the user has good knowledge of local forest conditions and reproduction patterns.

Specification of species

Species of the understory can be specified via two options. In the first option the rates of species composition can be specified that follow the database values used for model development in STAG. These rates are given in Table 10:

Table 10. Percentages of Species by Timber Type Rounded to the Nearest Five Percent[6].

	Timber Type											
	Douglas-fir			Mixed Conifer			Ponderosa Pine			True Fir		
Spp.	Diameter range			Diameter range			Diameter range			Diameter range		
	1-6*	6-11	1-11	1-6	6-11	1-11	1-6	6-11	1-11	1-6	6-11	1-11
PP	5	0	5	10	20	15	50	80	65	0	0	0
SP	5	0	5	5	5	5	5	0	5	5	5	5
IC	10	5	5	30	20	30	15	10	10	10	5	10
DF	60	85	70	15	20	15	5	5	5	0	0	0
WF	20	10	15	40	35	35	25	5	15	75	75	75
RF	0	0	0	0	0	0	0	0	0	10	15	10

We were unable to develop meaningful equations for prediction of species composition as related to the overstory composition and sizes of trees. Because of this the percentage of trees occurring in each species in the understory can be specified directly by the user of the program.

Creating an understory tree list

Because its possible to generate a large number of understory trees we use the following methodology to reduce the number of tree records being written. For either the predicted or the user specified number of understory trees (≤ 11.0 inches at DBH) we generate individual tree records with a tree expansion factor of one. The diameters of these trees are generated from the Weibull distribution using the equations for $\hat{b}$ and $\hat{c}$ given under "Generation of Understory trees." Tree heights and heights-to-crown base are determined according to equations [2] and [4] as described in "Estimating Total Height" and "Estimating Height To-Crown Base", respectively. Stochastic errors are added according to the methods described in "Stochastic Errors."

[6] DF_{1-6} denotes Douglas-fir timber type. The row entries corresponding to this column show the percent of species in the Douglas-fir timber type for trees within the 1-6 inch DBH class ($1.5 \leq DBH \leq 5.5$). Other columns show the percentage of species for a given timber type in the 6-11 inch DBH class, and the 1-11 inch DBH class. Values of less than 5 percent have been deleted, and the other categories within a column have been proportionally adjusted and rounded to the nearest 5 percent.

After the understory is developed in the above fashion, the tree records are added to the existing overstory stand description. If the total number of records exceeds the 500 tree record limit imposed by CACTOS, or if the total number of tree records exceeds some user specified limit (which may be greater or less than 500), the user is given the option of "compressing" the understory tree list. Understory tree record compression is carried out by averaging those tree records which have similar tree attributes, then replacing those individual tree records with their average values and an appropriate expansion factor.

The compression algorithm is implemented as follows. Individual tree records are grouped into DBH, total height, and live crown ratio classes by species. If the first grouping does not sufficiently reduce the number of understory tree records, successively coarser and coarser classes are examined until the number of understory tree records is less than or equal to the number desired.

The first grouping uses $\frac{1}{2}''$ DBH classes, five dynamically determined height classes, and five dynamically determined live crown ratio classes. The height and live crown ratio classes are dynamically determined in the sense that the data determine the class limits and class intervals for each live crown ratio class nested within height class, where each height class is nested within DBH class. Thus, the maximum and minimum heights for the smallest DBH class will generally be different from those in the largest DBH class. Similarly, the largest and smallest live crown ratios found in the smallest height class of the smallest DBH class will generally be different from the largest and smallest live crown ratios found in the largest height class of the smallest DBH class, *etc.* We felt that these nested classes would retain more of the "individuality" of each tree record than would non-nested classes.

If the first grouping fails to meet the desired number of understory tree records, the groupings are made successively coarser in the following manner. First, the number of dynamically determined height classes is reduced to three. If this grouping is unsuccessful, the number of live crown ratio classes is reduced from five to three also. Next, 1" DBH classes are tried, then two height classes, then 2 crown ratio classes, then 2" DBH classes. As a last resort, one tree record per understory species is attempted, though a compression of this severity is certainly not recommended if the number of generated understory tree records far exceeds the number of species.

CONVERTING STAND TABLE DATA TO AN INDIVIDUAL TREE LIST

A stand table contains the numbers of trees by diameter class (usually classes are between 1 and 2 inches) and species. It is a common method for obtaining field data, but obviously individual tree information is lost. Its main advantage is that since diameters only have to be crudely approximated there are substantial time savings in collecting information. The number of trees in the diameter classes can be thought of as a discrete approximation to a continuous diameter distribution.

The methodology used to convert stand table data into individual tree data to produce a facsimile stand description closely parallels the technique used for continuous data which was previously described under the section entitled "Generating stands from summary statistics". We assume that the distribution of grouped diameters given species [p(DBH | Species)] follows a Weibull distribution. The probability of a tree height falling into some discrete height class given its species and DBH class [p(H | Species, DBH)], and the probability of a tree crown falling into some discrete class given its species, DBH and height class [p(HCB | Species, DBH, H)] are both hypothesized to follow a normal distribution. These assumptions were tested using a Kolmogorov-Smirnov test and found to be acceptable. For further detail see Van Deusen (1984).

Diameter Distributions

We postulated that distribution of diameters across diameter classes followed a Weibull distribution, but within a given diameter class we assumed that trees followed a uniform distribution. If diameter classes are not wide then this is a plausible assumption. We tested this latter assumption on 50 1/4$^{\underline{th}}$ acre plots (see Van Deusen 1984) and found that the simplifying assumption of uniform distribution of diameters within a diameter class yielded results quite similar to that obtained with using Weibull distribution across classes when diameter classes were no larger than two inches.

Height Distributions

An average value for height and height-to-crown base is predicted from equations [1] and [3] by using the diameter class mean. The predicted average height is used to locate the centroid of the height distribution within a diameter class (see Figure 1). The variance of the distribution is then approximated using the variance of the regression of the height prediction equation. We estimate the proportion of trees to allocate to a specific height class within a diameter class by determining the percentage of the area under the curve for each height class. We call this process distributional apportionment because we allocate (apportion) the number of trees per diameter class over the height classes using this methodology.

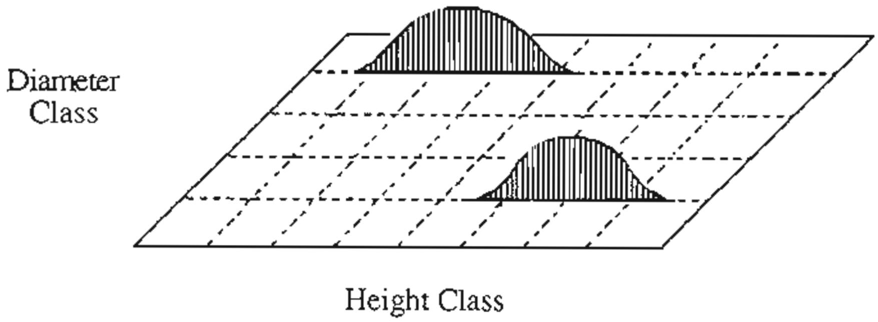


Figure 1. Distributional apportionment of stand table data.

Height-to-crown base Distributions

We assume that the distribution of crowns within a given height and DBH class follows a normal distribution. We allocate the numbers of trees into each of the crown classes using the same methodology as used for allocating trees into height classes. The normal curve is first located using the mean height-to-crown base value assuming the midpoints of the height and diameter class. The variance of the normal curve is approximated by the variance about the height-to-crown base predictive model. In the last step the area under the curve above each crown class is calculated and the number in each height-diameter cell over the crown class is determined according to these proportions.

This apportioning process calculates the numbers of trees to place in each cell of the height-diameter-crown categories. We define these cells to be either 1 or 2 inch diameter classes (specified by the user), 10 foot height classes, and 10 foot height-to-crown base classes. Individual tree dimensions (diameter, height, and height-to-crown base) are given an equal probability of occurring at any location within this three dimensional cell by drawing random numbers which correspond to x,y,z coordinates in 3D space. Using this procedure we have developed an individual tree list from the original stand table, but they are pseudo-individual in the sense that they have been estimated using the above procedure, rather than measured.

VALIDATING THE DIAMETER DISTRIBUTION GENERATION PROCEDURE

The procedure was tested on 166 one-fifth acre permanent plots from the Northern California Forest Yield Cooperative database described under DATA. Only the Weibull distribution was tested since the stands used for the test are considered to generally be managed stands and do not follow a negative exponential distribution. The second measurement data collected in 1984 from the southern Cascade region was used for the test. The accuracy of the procedure for predicting the number of trees per DBH class and the volume per DBH class was evaluated. Two inch DBH classes were used beginning at 5.5" and going to 49.5" DBH. The error index developed by Reynolds, Burk, and Huang (1989) was used for the test and is given as:

[28] $$e = N \cdot \sum_{j=1}^{k} \left| \int_{I_j} w(x) \cdot d\widehat{F(x)} - \int_{I_j} w(x) \cdot dF^*(x) \right|$$

where e = error index,
N = number of trees per acre,
$w(x)$ = weighting factor,
$\widehat{F(x)}$ = the cdf of diameters on a plot as predicted from the model,
$F^*(x)$ = the empirical cdf,
$dF(x)$ = the differential of the cdf (empirical or predicted) with respect to x (diameter)
k = the number of DBH classes,
I_j = the j^{th} DBH class.

As the authors of the index point out a "good" fit in one diameter class does not offset a "poor" fit in another. The error index provides a means for comparing the overall fit of a model to another model, but the individual cells (DBH, species classes) must be examined to determine where a particular model fits adequately.

We perform two sets of analysis. In the first we compute the error index of an "average stand". The "average stand" is the stand table produced by averaging all of the 166 stand tables associated with each of these plots. We judge our ability to produce a tree diameter distribution by seeing how accurately the number of trees in various diameter classes is predicted for this "average stand". We also judge how well our diameter distribution models work by comparing the volumes (Biging, 1983) predicted for each diameter class with the average volume computed from the 166 test plots. In the second analysis we present results which show the average of the error indices computed for each plot individually.

The results for the "average stand" are shown in Tables 11 and 12. Table 11 shows the "misclassification" by species and DBH class for the average of the 166 plots. By misclassification we mean the signed values calculated from differencing the predicted number of trees (or volumes) from the actual number of trees (or volumes) in each diameter class. The sum of the absolute values (predicted minus observed, see equation [28]) is used by Reynolds, and others to calculate the error index. Thus for ponderosa pine in the 8.5 DBH class this model underpredicted by 2 trees.

In the right margin of Table 11 are the error indices by species. The indices' magnitude correspond, relatively, to the abundance of the trees on the plots. In other words, the more trees there are the greater the error. The bottom margin is the average misclassification across species for a particular DBH class. Thus we see on average an underprediction for the 6.5 to 8.5 DBH classes and an overprediction in the 10.5 to 12.5 DBH classes. There are on average only slight underpredictions for the 18.5-22.5 and the 28.5 inch diameter class. The lower right cell of Table 11 provides the overall error index for this "averaged" plot which is a value of 68.

Another statistic we computed was the average plot error index with its associated standard errors. The average was 358 and the standard error was 11.8. The average is quite large and shows the difficulty of predicting the diameter distribution for a particular plot. The error index in Table 11 is much smaller (68) because we are averaging the plots and then computing the errors as opposed to average error index value (358.13) where we have the average of the individual plot error indices.

Table 12 provides the same type of information as Table 11. In Table 12 the error index is weighted by board foot volume whereas in Table 11 the error index was weighted by numbers of trees. In Table 12 we see that volumes are on average slightly underpredicted for the 6.5-8.5 inch diameter class. Volumes are overpredicted in the 10.5-18.5 inch diameter class, underpredicted in the 20.5-28.5 inch diameter class and overpredicted in the 30.5 inch and 32.5 inch and greater diameter class.

In Table 11 we reported that on average the models underpredicted by 2 or 3 trees in the 18.5-22.5 inch diameter class. Because trees in this size range average around 200-500 board feet it is not surprising that in Table 12 we find that the misclassification index for diameter classes in this range vary from an overprediction of 516 board feet to an underprediction of 2374 board feet. The average net effect of these over and underpredictions is a slight over prediction of 330 board feet. Thus there appears to be no major bias in volume associated with producing diameter distributions using the Weibull generation procedure.

We also computed the average plot error index weighted by volume with its associated standard errors. The average was 44540 board feet and the standard error was 3187 board feet. Again this underscores the difficulty of accurately predicting the diameter, and volume distribution on any particular plot.

In another test of this procedure we used the same plots to create "known" stand tables. We then used the stand tables to apportion the trees over height and crown classes. The results for the average stand table based on these 166 plots is presented in Table 13. The numbers of trees apportioned into these classes corresponded well with the actual numbers observed on the plots, except for the smallest diameter classes. Predicted heights and predicted heights-to-crown base were generally close to the observed average values. This demonstrates that stand tables can be generated which, on the average, closely approximate actual stands. Of course, good judgment should be exercised in using these routines. Real field data is always preferable to generating stands from summary statistics. Even though these procedures produce reasonable facsimiles to real stands there are always inaccuracies produced in this process. For a more detailed treatment of this analysis see Van Deusen (1984).

Table 11. Average misclassification indices (actual - predicted) of numbers of trees per acre by species and DBH classes and overall error index of the "average" plot from STAG version 4.0.

	Misclassification Index															
	DBH Class														Species	
Species	6.5	8.5	10.5	12.5	14.5	16.5	18.5	20.5	22.5	24.5	26.5	28.5	30.5	≥32.5	Error Index	
PP	3	2	-4	-6	0	0	0	1	1	0	1	0	0	0	-2	(18)
SP	1	0	0	0	1	0	0	0	0	0	0	0	0	0	2	(2)
IC	7	5	0	0	0	0	0	0	0	0	0	0	0	0	12	(12)
DF	5	0	-2	0	0	0	0	0	1	0	0	0	0	0	4	(8)
WF	9	4	-5	-5	-1	0	2	1	1	1	0	0	0	0	7	(28)
RF	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(0)
	25	11	-11	-11	0	0	2	2	3	0	1	0	0	0	22	(68)

Table 12. Average misclassification indices (actual - predicted) of board feet to a 6 inch top per acre by species and DBH classes from STAG version 4.0 The value in the last column is the signed value while next to it in () is the absolute value necessary for the computation of the error index.

	Misclassification Index															
	DBH Class														Species	
Species	6.5	8.5	10.5	12.5	14.5	16.5	18.5	20.5	22.5	24.5	26.5	28.5	30.5	≥32.5	Error Index	
PP	38	24	-221	-521	-345	-223	-295	318	671	186	983	-95	-288	-576	-344	(5394)
SP	2	-11	-6	-6	20	58	1	-110	60	125	-28	61	-67	-440	-341	(1651)
IC	7	-1	-51	-72	-35	-130	87	4	130	90	18	0	72	269	388	(988)
DF	21	-27	-193	-208	-215	-216	-299	-199	762	-3	199	0	-23	-161	-562	(2548)
WF	84	55	-369	-716	-652	-624	42	309	655	205	107	802	-38	526	386	(5236)
RF	3	-1	-44	-54	-57	-27	-52	18	96	72	384	-22	-69	-104	143	(1003)
	155	39	-884	-1577	-1284	-1162	-516	340	2374	675	1663	746	-413	-486	-330	(16820)

Table 13. Average Stand Table based upon 166 permanent plots from the Southern Cascade Region.

Diameter Class	Observed Av. Numbers	Expected Av. Numbers	Observed Av. Height	Expected Av. Height	Observed Av. HCB	Expected Av. HCB
6.5	50.2	21.7	35.5	39.7	21.0	16.3
8.5	49.7	35.9	45.2	48.4	26.1	21.3
10.5	30.4	42.5	56.4	57.6	31.9	26.3
12.5	28.9	40.7	65.0	65.5	36.3	30.5
14.5	29.5	29.7	72.2	73.4	38.5	34.6
16.5	18.8	19.8	80.8	81.0	42.7	38.3
18.5	15.2	12.5	85.7	87.9	44.7	42.4
20.5	10.3	8.0	95.0	95.0	49.0	42.5
22.5	9.8	4.5	99.1	102.3	51.6	49.0
24.5	4.6	3.0	106.6	109.2	52.7	51.8
26.5	5.0	1.9	113.1	114.8	59.4	55.2
28.5	3.0	1.5	116.3	120.0	61.3	58.1
30.5	1.3	1.0	120.0	124.8	65.8	60.7
32.5	0.6	0.5	119.6	132.1	61.7	65.8
34.5	0.7	0.4	135.2	136.6	75.5	69.9
36.5	0.6	0.3	121.9	149.3	64.6	71.3
38.5	0.4	0.2	140.4	145.8	69.6	68.5
40.5	0.1	0.1	135.2	157.7	76.6	73.6
42.5	0.1	0.1	147.7	168.8	89.7	78.1
44.5	0.0	0.0	0.0	170.3	0.0	78.1
46.5	0.1	0.1	146.3	167.8	97.0	83.1
48.5	0.0	0.0	0.0	169.8	0.0	85.3

DISCUSSION

The Stand Generator, STAG, is an important component of a simulation system for mixed conifer growth and yield projection. STAG was created to ensure that different types of inventory data could be supplemented to produce data sets suitable for projection in the forest simulator CACTOS. There are different procedures and analysis routines within STAG for 1) processing missing data, 2) converting stand table data (approximations to a diameter distribution), and 3) transforming summary statistics such as number of trees and basal area per acre to a stand description comprised of complete individual tree records for use in CACTOS. To "fill in" missing data STAG uses predictive equations for total height and height-to-crown base developed from a permanent plot system of over 20,000 trees in Northern California. To create a complete stand description based only on summary statistics (termed stand generation) is much more complicated. For this case STAG factors the joint distribution for species, DBH, H, and HCB into a product of probability density functions and models each of these components. The methodology developed for converting stand table data closely follows that described for stand generation.

These procedures are not intended to replace extensive data collection procedures. Instead these are intended to increase the availability of data that can be used with the CACTOS simulation system. The procedures developed for STAG have been tested using permanent plot data for mixed species, multiple aged coniferous stands. They produce relatively accurate and reliable results particularly when filling in missing data. The stand generation and stand table conversion techniques should be used more cautiously as they only produce a facsimile of a stand given the reduced data sets or summary statistics provided.

LITERATURE CITED

Avery, T. E., Burkhart, H. E. Forest Measurements. McGraw-Hill Book Co., New York, 331 pgs.

Biging, G.S. 1983. Volume tables for young-growth mixed conifers of Northern California based upon the stem analysis data. Part I: Volume tables for mixed conifers. Part II: Taper equations for mixed conifers. (Revised draft). Research Note #7. Northern Calif. Forest Yield Cooperative, Dept. of Forestry and Resource Mgmt., Univ. of Calif., Berkeley. 59 pgs.

Biging, Greg S. 1984. Taper equations for second-growth mixed conifers of northern California. For. Sci. 30(4): 1103-17.

Biging, G. S. and W. J. Meerschaert. 1990. STAG user's guide: The Forest Stand Generator for Mixed Conifer Species in California. Version 3.3. Res. Note. 21 revised. Northern California Forest Yield Cooperative. University of California, Berkeley. 36 pgs.

Biging, Greg S. and Lee C. Wensel. 1987. STAG: A forest STAnd Generator for producing complete CACTOS stand descriptions. In: Forest Growth Modelling and Prediction (A. Ek and T. Burk , eds.). USDA For. Serv. Gen. Tech. Rep. NC-120. Vol. 1, pgs. 47-53.

Biging, G. S., Meerschaert, W. J., and T. A. Robards. 1990. STAG user's guide: The Forest Stand Generator for Mixed Conifer Species in California. Version 3.3. Res. Note. 21. Northern California Forest Yield Cooperative. University of California, Berkeley. Draft ms.

Biging, Greg S., Meerschaert, Walter J., and Timothy A. Robards. 1991. STAG User's Guide: The Forest Stand Generator for mixed conifer species in California. Version 4.0. Manuscript for the Univ. of Ca. Ag. Exp. Stn. Bulletin series.

Burk T. E., M. H. Hansen and A. R. Ek. 1982. Combining sources of information for improved in-place inventory statistics. In Place Inventories, Orno ME. 1981.

Davis, L. S., and K. N. Johnson. 1987. *Forest Management*. 3rd Edition. McGraw-Hill. New York. 790 pgs.

Ek, Alan R. 1974. Nonlinear models for stand table projection in northern hardwood stands. Can. J. For. Res. 4:23-27.

Husch, B., Miller, C. I., and T.W. Beers. 1982. *Forest Mensuration*. 3rd Edition. John Wiley & Sons. New York. 402 pgs.

Hyink, D. M. and J. W. Moser. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. For. Sci. 29:85-95.

Knoebel, B.R. and Burkhart, H.E. 1991. A bivariate distribution approach to modeling forest diameter distributions at two points in time. Biometrics 47(1):241-253.

Knox, R. G., Peet, R. K., and N. L . Christensen. 1989. Population dynamics in loblolly pine stands: changes in skewness and size inequality. Ecology 70: 1153-1166.

Meerschaert, Walter J. 1987. An overview of CACTOS: the California Conifer Timber Output Simulator. Proceedings, IUFRO Forest Growth Modelling and Prediction Conference, Minneapolis, MN, August 24-28, 1987.

Meyer, H. A. 1952. Structure, growth, and drain in balanced uneven-aged forests. J. For. 50(2):85-92.

Miller, T. E. and J. Weiner. 1989. Local density variation may mimic effects of asymmetric competition on plant size variability. Ecology 70: 1188-1191.

Reineke, L.H. 1933. Perfecting a stand density index for even-aged forest. J. Agric. Res. 46:627-638.

Reynolds, Marion R., T.E. Burk, and W.C. Huang. 1988. Goodness of fit tests and model selection procedures for diameter distribution models. For. Sci. 34(2):373-399.

Shifley, S. R. 1990. Analysis and modeling of patterns of forest ingrowth in the North Central United States. Ph.D. dissertation, Univ. of Minnesota, St. Paul, MN.

Shifley, S. R. and E. Lentz. 1985. Quick estimation of the three-parameter Weibull to describe tree size distributions. Forest Ecology and Mgt. 13: 195-203.

Van Deusen, Paul C. 1984. Complete and partial generation of tree characteristics for mixed species stands. Ph.D. dissertation, Univ, of California, Berkeley. 122 pages.

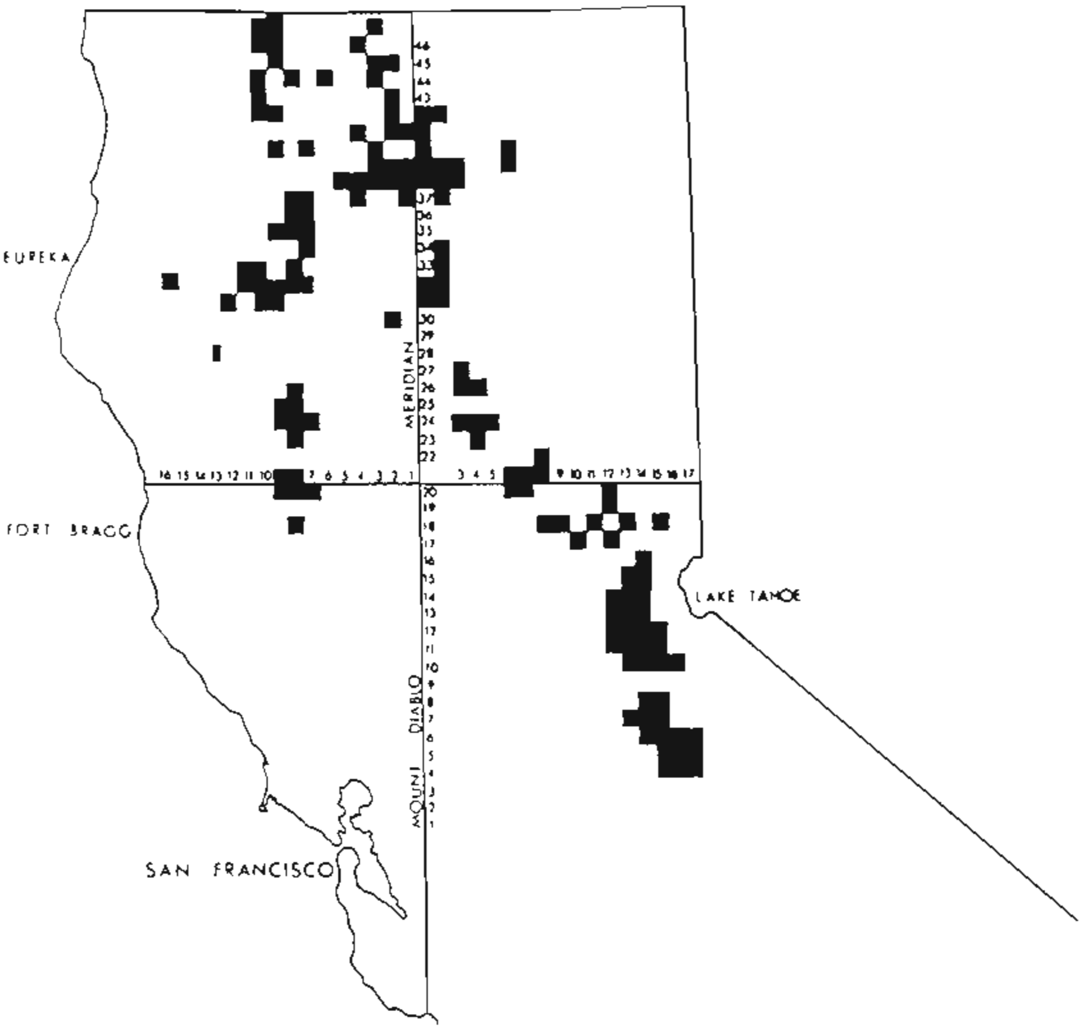
Van Deusen, Paul C. and Greg S. Biging. 1985. STAG: A STAnd Generator for Mixed Species Stands. Version 2.0. Research Note No. 11. Dept. Forestry and Res. Mgt. Univ. Calif. Berkeley. 25 pp.

Wensel, L. C., P. J. Daugherty, and W. J. Meerschaert. 1986. CACTOS user's guide: The California conifer timber output simulator. Div. of Ag. Sci. Univ. Calif., Berkeley. Bulletin 1920. 91p.

Wensel, L. C., W. J. Meerschaert, and G. S. Biging. 1987. Tree height and diameter growth models for northern California conifers. Hilgardia 55(8) 1-20.

Wensel, Lee C. and Greg S. Biging. 1987. The CACTOS System for individual-tree growth simulation in the mixed conifer forests of California. *In: Forest Growth Modelling and Prediction.* (A. Ek and T. Burk, eds.). USDA For. Serv. Gen. Tech. Rep. NC-120. Vol. 1, pgs. 175-183.

Figure 1. Location of permanent plots by township.



APPENDIX

Summary Statistics for the permanent plot tree data.

Ponderosa Pine				n = 4173
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	14.11	6.58	5.5	55.8
Total Height (ft)	72.87	27.97	12.0	184.0
Height-to-crown base (ft)	34.92	18.21	1.0	145.0
SITE	74.20	17.51	29.0	150.0
Basal Area (ft^2) per acre	104.42	82.31	22.5	532.7
Number of trees per acre	215.43	100.04	16.0	515.0

Sugar Pine				n = 1070
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	15.96	8.41	5.5	59.1
Total Height (ft)	73.49	31.46	15.0	199.0
Height-to-crown base (ft)	36.91	18.25	1.0	105.0
SITE	76.67	16.30	29.0	150.0
Basal Area (ft^2) per acre	216.27	98.08	30.2	532.7
Number of trees per acre	207.89	90.64	15.0	490.0

Incense Cedar				n = 2260
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	12.70	6.90	5.5	67.6
Total Height (ft)	48.19	22.17	11.0	182.0
Height-to-crown base (ft)	25.10	14.88	1.0	95.0
SITE	76.22	16.49	29.0	130.0
Basal Area (ft^2) per acre	213.25	90.05	27.2	532.7
Number of trees per acre	205.21	91.31	18.0	515.0

Douglas-fir				n = 2458
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	13.22	5.91	5.5	50.4
Total Height (ft)	72.59	24.92	11.0	170.0
Height-to-crown base (ft)	38.63	18.34	2.0	126.0
SITE	77.82	16.59	36.0	157.0
Basal Area (ft^2) per acre	169.06	72.56	16.5	424.2
Number of trees per acre	182.98	76.14	20.0	515.0

White fir				n = 5167
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	13.03	6.17	5.5	48.9
Total Height (ft)	64.29	26.60	9.0	171.0
Height-to-crown base (ft)	33.37	17.49	1.0	114.0
SITE	76.29	16.50	23.0	130.0
Basal Area (ft^2) per acre	211.82	89.90	3.5	532.7
Number of trees per acre	203.20	89.92	5.0	525.0

Red fir				n = 501
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	15.39	7.54	5.5	51.6
Total Height (ft)	70.01	29.09	15.0	154.0
Height-to-crown base (ft)	33.55	18.63	3.0	92.0
SITE	65.43	11.35	46.0	104.0
Basal Area (ft^2) per acre	230.00	94.37	36.6	428.8
Number of trees per acre	199.41	95.00	15.0	525.0

Other Hardwoods				n = 273
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	10.79	4.41	22.5	454.4
Total Height (ft)	47.57	19.85	11.0	104.0
Height-to-crown base (ft)	25.36	11.62	2.0	69.0
SITE	78.51	21.52	37.0	114.0
Basal Area (ft^2) per acre	204.61	108.70	22.5	454.4
Number of trees per acre	173.97	65.70	36.0	305.0

Black Oak				n = 340
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	12.93	7.00	5.5	52.7
Total Height (ft)	52.55	19.55	12.0	164.0
Height-to-crown base (ft)	24.75	13.92	1.0	82.0
SITE	75.83	15.45	37.0	114.0
Basal Area (ft^2) per acre	188.40	86.57	30.6	424.2
Number of trees per acre	184.35	79.71	16.0	376.0

APPENDIX B

Summary Statistics for the permanent plot small tree data.

All species		Tree Statistics		n = 3339
Variable	Mean	Std. Dev.	Minimum	Maximum
DBH (in)	3.51	1.16	1.50	5.40
Total Height (ft)	19.24	8.50	5.00	64.00
Height-to-crown base (ft)	10.34	7.30	1.00	56.00
	Plot Statistics			n=308
Basal Area of all tress ≤ 5.5"	10.51	8.78	0.59	51.07
Number of all trees ≤ 5.5"	89.93	122.85	4.00	755.00