Estimation of Relative Density for Douglas-fir Stands

Mark Senger
Econometrics 482 –Prof. Startz
June 5, 2002

EXECUTIVE SUMMARY

There has been extensive forestry research examining the effect of density on the growth of trees in plantations. Density can be expressed as trees per unit area, but tree size is also an important component of density because trees increase in size during their lifespan. A number of density indexes have been developed to quantify the relationship, and provide a basis for making comparisons between different locations. These indexes relate the current condition to a theoretical maximum density which can be estimated, and is unique for a given tree species. Measures of density are used by forest managers to make assessments of the condition of a forested area, and decisions about activities such as tree harvesting.

A regression model was developed that relates a density index to other tree attributes that are more easily assessed in the field. The data used for this analysis are from permanent experimental plots that have been established by a research organization called the Stand Management Cooperative. The chosen variables were live crown ratio (plot_lcr), breast height age (bhage), dominant tree height (ht40) and trees per acre (stems_acre). These tree attributes were shown to be correlated to relative density (RD), and various functional forms were tested to express the relationships between the underlying variables. The method of Weighted Least Squares was used to correct for observed heteroskedasticity. All of the coefficients in the fitted equation shown below were highly significant, and the regression had an R^2 of 0.89 and an SER of 3.86.

RD= -20.51 + 6.66(plot_lcr) + 0.38(ht40) + 0.14(stems_acre) + 0.53(plot_lcr*ht40) -0.73(plot_lcr*bhage) - 0.12(plot_lcr*stems_acre)

The model provides a reasonable representation of reality, and forestry literature and theory support the relationships that are observed between the variables in the fitted equation. The utility of this regression equation may be as a practical tool that reduces the time and money required for the measurement of trees in the field.

Given the limited resources (light, water, nutrients, etc.) defined as "growing space" that are available on a site, the amount of growing space available to an individual tree is determined by the number and size of adjacent trees. A group of trees (referred to as a "stand") is the level at which density is a relevant concept, although density does affect individual tree growth. Theories of tree growth also assume that while total growth per unit area will be relatively constant within some range of tree density, this growth will be distributed as progressively less growth per tree as tree size and/or the number of trees per unit area increases. There has been extensive forestry research examining the relationship between density and tree growth, and a number of density indexes have been developed to quantify the relationship. These indexes measure the "occupied" growing space compared to a theoretical maximum density which can be estimated, and is unique for a given tree species. Trees on a given site compete for resources, so as a stand approaches maximum density there is a loss of tree growth and mortality. Forest managers often use a relative density index to decide when to remove some of the trees (referred to as thinning) to maintain the growth and health of trees remaining on a site. The utility of these indexes is that they are independent of site quality (soil productivity differences) and tree age, allowing them to be applied in a wide variety of conditions. One of these indexes is Curtis' relative density (RD)[1], which is easier to compute than others because it is based on diameter. Although tree diameter is easy to measure, extensive sampling and calculations are required to obtain the relative density of a stand.

THE MOTIVATION

With this regression model I am attempting to develop a practical tool that expresses relative density as a function of tree attributes that are more easily estimated "by eye" in the field. Even experienced foresters have difficultly giving an accurate estimate of relative density based on what they see in the field, and relative density must be calculated at a later time in the office. While diameter is an easy tree attribute to measure, obtaining a sample to calculate relative density can take hours of field work. Tree (or stand) characteristics such as age and number of trees per acre are known for planted stands, and other tree attributes can be quickly measured or estimated with reasonable accuracy "by eye" while in the field. An equation that establishes a meaningful relationship between relative density and other tree attributes could

[1] See Appendix A for Curtis' Relative Density Formula

be used to determine whether more extensive measurements are needed at a location. As a method of quickly assessing relative density, the equation would save time and money by avoiding unnecessary measurements when the stand is not near a target level of relative density.

DATA

The dataset that I used for this project comes from the Stand Management Cooperative, a research organization that has established permanent experimental plots throughout the Northwest. These plots are used to test the effects of various treatments on tree growth and stand development, and density is one of these treatments. Relative density is a per unit area concept, and so I used data which are the mean plot values (per unit area) for given tree attributes. I limited the data to plots of Douglas-fir trees, because tree growth patterns, relationships between given tree attributes, and theoretical maximum density all vary by tree species. I further limited the data to stands which had a quadratic mean diameter (QMD) greater than 2.0 inches, because the literature shows that Curtis' relative density is not applicable to younger stands (QMD<2). For some of the plots there was no data for tree age, and I excluded these plots because I wished to use age as an independent variable. From the original dataset of 2055 plots, these criteria resulted in a sample of 1239 plots of Douglas-fir trees that were of sufficient size for my analysis.

The independent variables that I used in my regression are live crown ratio (plot_lcr), dominant tree height (ht40), breast height age (bhage), and trees per acre (stems_acre). Live crown ratio is the proportion of a tree that has living branches, the "factory" which produces resources necessary for tree growth, and plot_lcr is the average live crown ratio for a given plot. Dominant tree height (ht40) is an average height for the 40 tallest trees on a site, and provides a measure of tree size. Breast height age (bhage) is the tree age where tree diameter is measured, which is 4.5 feet above the ground. Trees per acre (stems_acre) is implicit in the relative density equation, however two sites with the same relative density can result from different combinations of tree per acre and tree diameter, so I felt including this factor would account for some of the resulting variability.

THE MODEL

In developing the regression model, I first considered how each of my independent variables was related to relative density. A correlation matrix created in EVIEWS revealed that each of the four independent variables was correlated rather strongly with relative density. I plotted each of the chosen variables against relative density as an aid in determining which functional form would be most appropriate for each factor. The graphs showed that age and height had strong linear relationships with relative density, but exhibited increasing variation with increasing relative density (heteroskedasticity). Stems per acre showed a weaker linear relationship with relative density. The relationship between live crown ratio and relative density appeared to be curvilinear, and in light of this relationship I chose to include a quadratic factor for live crown ratio. After re-examining the correlation matrix, I included interaction terms for live crown ratio with each of the other factors, because live crown ratio is correlated with each of them, and I felt that scaling each of the other variables with live crown ratio could provide a better representation of the true model. I first ran an OLS regression on the preliminary model below:

$RD = B_0 + B_1$ (plot_lcr) + B_2 (plot_lcr^2) + B_3 (ht40) + B_4 (bhage) + B_5 (stems_acre) + B_6 (plot_lcr*ht40) + B_7 (plot_lcr*bhage) + B8 (plot_lcr*stems_acre)

In EVIEWS, I used White's Test for Heteroskedasticity to determine if this test would detect what I had observed in the plots of relative density against bhage and ht40. The results of the test showed that heteroskedasticity was present, so a Weighted Least Sqauares (WLS) regression would be needed to get accurate standard errors and t-statistics.

I used a method discussed in lecture to "correct" for the heteroskedacticity. I took the residuals from the OLS regression of the preliminary model (as an estimate of the true errors), and regressed these squared residuals on a constant, bhage and ht40 (the two terms for which I observed heteroskedasticity with my original bivariate plots). I then used the fitted values from this auxiliary regression as an estimate of the variance of the error. Following EVIEWS convention, I used the inverse of the standard error as the weight for the WLS regression of the preliminary model.

A comparison of the OLS and the WLS regressions (OUTPUTS 1 & 2)[2] reveals differences in the estimated coefficients. In general, the estimated coefficients given by the two methods are similar as was expected (they should both be unbiased estimates of the betas), with the notable exceptions of the live crown ratio quadratic term and breast height age. Age was not significant in the OLS model (bearing in mind the t-statistics are not accurate here), and became even less significant (p=0.7952) after weighting the regression. The coefficient for the squared live crown ratio term lost significance in the WLS model (p= 0.1049), which may mean that the observed curvilinear pattern was a result of heteroskedasticity rather than functional form. I decided to omit these two terms from the final model, because neither was significant at the 95% level. The WLS regression (Output 2) [3] revealed that the coefficients for all of the other terms were significant at the 95% level. To obtain weighting values for the new model, I applied the same method described previously to the new equation (without bhage and plot_lcr2). I then used these new weights for a WLS regression on the final model. The final model and fitted equation are shown below.

RD= B_0 +B_1 (plot_lcr) + B_2 (ht40) + B_3 (stems_acre) + B_4 (plot_lcr*ht40) + B_5 (plot_lcr*bhage) + B_6 (plot_lcr*stems_acre)

The fitted equation:

RD= -20.51 + 6.66(plot_lcr) + 0.38(ht40) + 0.14(stems_acre) + 0.53(plot_lcr*ht40) -0.73(plot_lcr*bhage) - 0.12(plot_lcr*stems_acre)

Analysis of the final model reveals that it follows theoretical relationships between relative density and the tree attributes that were used in the regression. The fitted equation provides a reasonable fit for the data, with an R^2 of 89%, and a SER value of 3.86 (Output 3)[4]. The negative value for the intercept in this model has no significance, because with the restriction of QMD>2 all of the right-hand side variables will have values greater than zero. The positive coefficients for dominant height (ht40) and stems per acre were expected, showing an increase in relative density with an increase in tree “size” and number of trees per unit area respectively. The positive coefficient for live crown ratio (plot_lcr) term suggests that relative density goes down as the live crown ratio is reduced, which is the opposite effect that was expected. However, the partial effect of live crown ratio also includes the interaction terms which may correct this observation, depending on the values of bhage, ht40 and stems_acre. Several of the interaction terms may seem to be redundant, but together they provide a clearer picture of the effect of the live crown on relative density. The

---

[2] See Appendix B

[3] See Appendix B

[4] See Appendix B

interaction of dominant height and live crown ratio is simply the length in feet of the live crown on the dominant trees, and as expected, this coefficient has a positive value. The coefficient for the interaction of live crown and breast height age is expected to be negative, because with greater age the live crown ratio tends to go down as each tree competes more with neighboring trees. The interaction between live crown ratio and stems per acre gives a rough indication of how much total foliage is on an acre, by applying the average foliage per tree across all trees that are present. Theories of tree growth indicate that above some minimum density, the amount of foliage per unit area is relatively constant, with more or less foliage allocated to each tree as the number of trees per unit area varies. The negative coefficient suggests that an increase in the amount of foliage results in a lower relative density, which seems to be counterintuitive, so this term may be "adjusting" the main effects of the live crown ratio and stems per acre terms to capture variation in one or both terms.

CONCLUSIONS

The final regression model provides a relatively good fit for the data that were used in this analysis, and may work well as a diagnostic tool. With a standard error of the regression of 3.86, the equation should produce estimates that are more accurate than estimates done “by eye” in the field (with standard errors of 10-15). A prediction from the model with a confidence interval that includes the "target" relative density for a given stand would result in the need to perform more accurate measurements and compute the actual relative density. By reducing the number of stands in which these relative density measurements need to be implemented, the model would save time and money.

One potential problem with the model is the use of trees per acre as an independent variable. In constructing the model, I made the assumption that this parameter would be known without taking measurements. It is true that the trees per acre will be known for a previous time period, but if there has been sufficient mortality in a stand the prediction of relative density would be inaccurate (dead trees are not included in relative density). There are two possible solutions to this oversight. The first is to use the model unless a substantial number of dead trees are observed in the stand. The second alternative is to take a sample to determine trees per acre, which would be less time consuming than the sample needed for relative density calculations.

APPENDIX A

Curtis' Relative Density (RD) is defined as: $RD = BA*QMD^{-0.5}$

Where quadratic mean diameter (QMD) is: $QMD = (S\ (d_i^2)/n)^{0.5}$

And Basal Area (BA) expressed in square feet per acre is:

$BA = (p*QMD^2*N)/(4*144)$ or the equivalent $BA = k*N*QMD^2$

With the variables defined as:

d_i = tree diameter at 4.5 feet above the ground measured in inches
n = number of trees measured
k = a constant (0.005454 for English units)
N = number of trees per acre

APPENDIX B

Dependent Variable: RD
Method: Least Squares
Date: 05/14/02 Time: 11:35
Sample(adjusted): 1 1323 IF INSTALL<900 AND QMD>2 AND BHAGE>0
Included observations: 1239 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-21.85212	6.258520	-3.491580	0.0005
PLOT_LCR	25.48989	14.34504	1.776913	0.0758
PLOT_LCR2	-18.25193	8.071865	-2.261179	0.0239
HT40	0.254641	0.076811	3.315158	0.0009
BHAGE	0.280546	0.190869	1.469835	0.1419
STEMS_ACRE	0.136811	0.004026	33.98031	0.0000
PLOT_LCR*HT40	0.644569	0.112665	5.721108	0.0000
PLOT_LCR*BHAGE	-1.100102	0.288147	-3.817852	0.0001
PLOT_LCR*STEMS_ACRE	-0.115869	0.005000	-23.17192	0.0000

R-squared	0.933863	Mean dependent var	27.91256
Adjusted R-squared	0.933433	S.D. dependent var	16.62988
S.E. of regression	4.290604	Akaike info criterion	5.757969
Sum squared resid	22643.42	Schwarz criterion	5.795176
Log likelihood	-3558.062	F-statistic	2170.978
Durbin-Watson stat	0.866712	Prob(F-statistic)	0.000000

OUTPUT 1

Dependent Variable: RD
Method: Least Squares
Date: 06/04/02 Time: 11:17
Sample(adjusted): 1 1323 IF INSTALL<900 AND QMD>2 AND BHAGE>0
Included observations: 1239 after adjusting endpoints
Weighting series: W

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-29.19046	5.902867	-4.945133	0.0000
PLOT_LCR	27.49400	13.08162	2.101728	0.0358
PLOT_LCR2	-11.83190	7.291164	-1.622773	0.1049
HT40	0.416976	0.081202	5.135027	0.0000
BHAGE	0.053094	0.204521	0.259599	0.7952
STEMS_ACRE	0.144184	0.003926	36.72752	0.0000
PLOT_LCR*HT40	0.482739	0.108601	4.445051	0.0000
PLOT_LCR*BHAGE	-0.808226	0.284007	-2.845796	0.0045
PLOT_LCR*STEMS_ACRE	-0.126070	0.004622	-27.27431	0.0000

Weighted Statistics

R-squared	0.892188	Mean dependent var	24.51705
Adjusted R-squared	0.891487	S.D. dependent var	11.70970
S.E. of regression	3.857324	Akaike info criterion	5.545062
Sum squared resid	18301.10	Schwarz criterion	5.582268
Log likelihood	-3426.166	F-statistic	2574.490
Durbin-Watson stat	0.909579	Prob(F-statistic)	0.000000

Unweighted Statistics

R-squared	0.932222	Mean dependent var	27.91256
Adjusted R-squared	0.931781	S.D. dependent var	16.62988
S.E. of regression	4.343525	Sum squared resid	23205.43
Durbin-Watson stat	0.852054		

OUTPUT 2

Dependent Variable: RD
Method: Least Squares
Date: 06/03/02 Time: 11:27
Sample(adjusted): 1 1323 IF INSTALL<900 AND QMD>2 AND BHAGE>0
Included observations: 1239 after adjusting endpoints
Weighting series: LAST_WGHT

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-20.51412	2.557314	-8.021745	0.0000
PLOT_LCR	6.661044	2.747393	2.424496	0.0155
HT40	0.381392	0.028575	13.34725	0.0000
STEMS_ACRE	0.143422	0.003893	36.84017	0.0000
PLOT_LCR*HT40	0.531458	0.040065	13.26474	0.0000
PLOT_LCR*BHAGE	-0.729591	0.068138	-10.70756	0.0000
PLOT_LCR*STEMS_ACRE	-0.124951	0.004573	-27.32186	0.0000

Weighted Statistics

R-squared	0.892453	Mean dependent var	24.54194
Adjusted R-squared	0.891929	S.D. dependent var	11.74211
S.E. of regression	3.860121	Akaike info criterion	5.544908
Sum squared resid	18357.46	Schwarz criterion	5.573846
Log likelihood	-3428.070	F-statistic	3425.755
Durbin-Watson stat	0.914855	Prob(F-statistic)	0.000000

Unweighted Statistics

R-squared	0.932037	Mean dependent var	27.91256
Adjusted R-squared	0.931706	S.D. dependent var	16.62988
S.E. of regression	4.345908	Sum squared resid	23268.68
Durbin-Watson stat	0.858249		

OUTPUT 3