by David N. Larsen
and David A. Gansner

# SIMPLE MODELS FOR ESTIMATING LOCAL REMOVALS OF TIMBER IN THE NORTHEAST

USDA FOREST SERVICE RESEARCH PAPER NE-319
1975

NORTHEASTERN FOREST EXPERIMENT STATION
FOREST SERVICE, U.S. DEPARTMENT OF AGRICULTURE
6816 MARKET STREET, UPPER DARBY, PA. 19082
F. BRYAN CLARK, STATION DIRECTOR

**THE AUTHORS**

DAVID N. LARSEN received his bachelor of science degree in forestry from Washington State University in 1964, a master of science degree in forest economics from Michigan State University in 1967, and a master of arts degree in regional science from the University of Pennsylvania in 1971. He joined the Northeastern Forest Experiment Station at Upper Darby, Pennsylvania, in 1967 and is a research forester in the regional production economics work unit.

DAVID A. GANSNER is leader of the Northeastern Forest Experiment Station's regional production economics research work unit at Upper Darby, Pennsylvania. He received his degree in forestry from the University of Missouri in 1958, a master's degree in agricultural economics from Ohio State University in 1968, and a master's degree in regional science from the University of Pennsylvania in 1970. He has been with the USDA Forest Service since 1958.

MANUSCRIPT RECEIVED FOR PUBLICATION 28 JUNE 1974

# SIMPLE MODELS FOR ESTIMATING LOCAL REMOVALS OF TIMBER IN THE NORTHEAST


## ABSTRACT

Provides a practical method of estimating subregional removals of timber and demonstrates its application to a typical problem. Stepwise multiple regression analysis is used to develop equations for estimating removals of softwood, hardwood, and all timber from selected characteristics of socioeconomic structure.


## INTRODUCTION

Resource analysts at the Northeastern Forest Experiment Station are frequently asked (by resource planners, wood-industry representatives and others) to provide up-to-date estimates of local removals of timber in the Northeast. The Station's Forest Survey work unit helps meet this need by publishing annual reports on the production of pulpwood and data on all timber removals in conjunction with their periodic statewide forest inventories. But the timber cut for pulpwood is only a portion of the total removed, and timber-removal data collected at the time of state inventories are not current for many areas and are often not available for particular localities of interest. To solve this problem, we have developed models for estimating local timber removals. The models are easy to use and appear to work quite well for subregions of a million acres or more.

## APPROACH

Underlying the development of the models was the premise that timber removal in any particular locale is related to, and can therefore be expressed as a function of, that locale's socioeconomic structure. Stepwise multiple regression analysis was used to develop equations for estimating timber removals as a function of selected characteristics of socioeconomic structure. Recent information compiled for 47 sample subregions in 12 northeastern states, (Connecticut, Delaware, Kentucky, Maine, Massachusetts, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Vermont and West Virginia) provided data for the analysis.

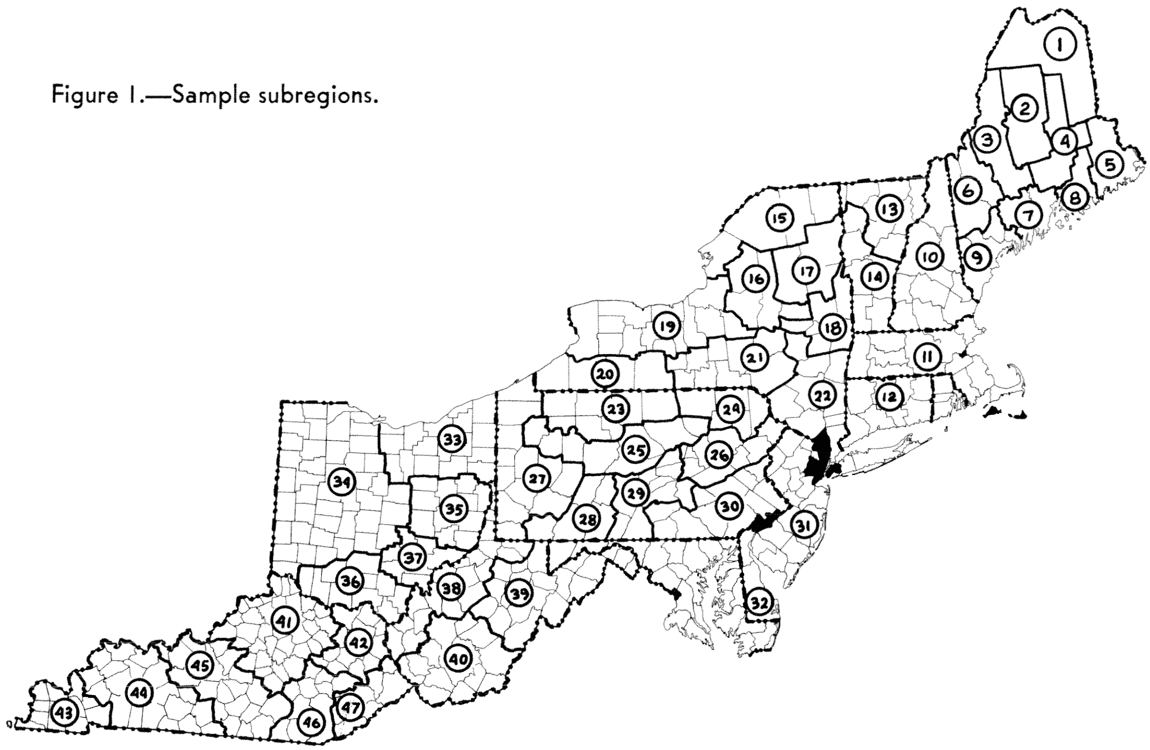

Figure 1.—Sample subregions.

Locations of the subregions are shown in figure 1.

The size of the sample subregions varies considerably, from 1 million to 12 million acres. We wanted to reduce the effect of subregion size as much as possible, so all variables whose values might be influenced by subregion size were expressed in units per acre of total land area.

## Regressand Variables

Three regressand variables were selected for analysis:

$Y_A$ = Current annual removal of *all timber* in cubic feet per acre of total land,

$Y_S$ = Current annual removal of *softwood timber* (conifers: spruce, fir, pine, hemlock, etc.) in cubic feet per acre of total land, and

$Y_H$ = Current annual removal of *hardwood timber* (broad-leaved trees: oak, yellow poplar, maple, ash, etc.) in cubic feet per acre of total land.

Estimates of current net annual removals of growing stock (for all purposes: wood products, land clearing, etc.) for each sample subregion were derived from records of the Northeastern Station's Forest Survey work unit.

## Regressor Variables

Sixteen regressor variables were tested and the five marked by asterisks proved to be significant:

$X_1$ = Proportion of commercial forest land in public ownership

* $X_2$ = Softwood growing stock inventory in thousands of cubic feet per acre of total land

* $X_3$ = Hardwood growing stock inventory in thousands of cubic feet per acre of total land

* $X_4$ = Softwood sawtimber volume inventory in thousands of board feet per acre of total land

$X_5$ =Hardwood sawtimber volume inventory in thousands of board feet per acre of total land

$X_6$ =Softwoods as a proportion of total growing stock

$X_7$ =Commercial forest land as a proportion of total land

* $X_8$ =Number of lumber and wood-product firms (Standard Industrial Classification [SIC] 24) per thousand acres of total land

$X_9$ =Value added by all manufacture per thousand acres of total land

$X_{10}$=Furniture and lumber and wood-product employees (SIC 24 and 25) per thousand acres of total land

$X_{11}$=Furniture and lumber and wood-product employees (SIC 24 and 25) as a proportion of total employees

$X_{12}$=Total human population per acre of total land

$X_{13}$=Rural human population per acre of total land

$X_{14}$=Average level of education for males 25 years of age and more

* $X_{15}$=Per capita income in thousands of dollars

$X_{16}$=Rural road mileage per thousand acres of total land

There were two major reasons for selecting this particular set of variables. First, we feel they include the socioeconomic conditions that determine the amount of timber removed in any subregion of the Northeast. Second, data on each of these variables are readily available, for each county of the Northeast, from Census and Forest Survey publications.

Data for these variables provided the input for the regression equations. The coefficient for each successive regressor variable had to be significant at the .01 level to be included in an equation.

## RESULTS

Stepwise multiple regression analyses gave us three simple equations for estimating current annual removals of timber (per acre of total land area) for subregions in the Northeast:

(1) All timber;

$$Y_A = 15.13 + 13.70X_2 + 116.49X_8 - 4.78X_{15} - 2.25X_4$$

$$R^2 = 0.84$$

(2) Softwood timber;

$$Y_S = -1.29 + 17.02X_2 + 28.99X_8$$

$$R^2 = 0.90$$

(3) Hardwood timber;

$$Y_H = 9.55 + 5.53X_3 + 63.01X_8 - 2.99X_{15} - 1.93X_4$$

$$R^2 = 0.69$$

Five of the original 16 regressor variables are included in these equations and their inclusion makes good sense. Variables $X_2$ (inventory of softwood growing stock volume), $X_3$ (inventory of hardwood growing stock volume), and $X_4$ (inventory of softwood sawtimber volume), are important indicators of the availability of local timber supplies. Variable $X_8$ (number of lumber and wood product firms) reflects local levels of market activity and timber demand. Variable $X_{15}$ (per capita income) reflects affluence and the degree to which other land uses compete with timber production.

The position of the regressor variables in each equation indicates the relative importance of each variable as an estimator of timber removal. Variables on the left are more important than those on the right. For example, $X_2$ is much more important than $X_4$ in equation (1).

The coefficient of multiple determination ($R^2$) for each equation is at least 0.69. Thus the rather small number of regressor variables included in these equations explain a rather large share of the variation in timber removal for the sample subregions; the models fit the sample data quite well.

## Table 1. Comparison of equation estimates with recorded estimates of current annual removals per acre of total land

| Subregion | All timber removals: Equation | All timber removals: Recorded | All timber removals: Difference | All timber removals: Relative error | Softwood timber removals: Equation | Softwood timber removals: Recorded | Softwood timber removals: Difference | Softwood timber removals: Relative error | Hardwood timber removals: Equation | Hardwood timber removals: Recorded | Hardwood timber removals: Difference | Hardwood timber removals: Relative error |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cubic Feet | | | Percent | Cubic Feet | | | Percent | Cubic Feet | | | Percent |
| 1 | 18.96 | 19.29 | —0.33 | 1.7 | 15.16 | 16.27 | —1.11 | 6.8 | 4.83 | 3.03 | 1.80 | 59.4 |
| 2 | 20.58 | 20.32 | .26 | 1.3 | 16.84 | 14.76 | 2.08 | 14.1 | 5.76 | 5.56 | .20 | 3.6 |
| 3 | 19.28 | 17.92 | 1.36 | 7.6 | 13.90 | 13.47 | .43 | 3.2 | 6.56 | 4.45 | 2.11 | 47.4 |
| 4 | 21.76 | 26.59 | —4.83 | 18.2 | 13.56 | 16.99 | —3.43 | 20.2 | 8.20 | 9.60 | —1.40 | 14.6 |
| 5 | 18.08 | 23.76 | —5.68 | 23.9 | 12.05 | 16.16 | —4.11 | 25.4 | 6.17 | 7.60 | —1.43 | 18.8 |
| 6 | 23.22 | 25.21 | —1.99 | 7.9 | 10.98 | 9.53 | 1.45 | 15.2 | 11.40 | 15.68 | —4.28 | 27.3 |
| 7 | 16.47 | 20.02 | —3.55 | 17.7 | 8.20 | 12.55 | —4.35 | 34.7 | 6.98 | 7.46 | — .48 | 6.4 |
| 8 | 16.73 | 11.12 | 5.61 | 50.4 | 11.20 | 8.67 | 2.53 | 29.2 | 5.50 | 2.44 | 3.06 | a |
| 9 | 23.54 | 17.52 | 6.02 | 34.4 | 10.78 | 11.53 | — .75 | 6.5 | 10.46 | 5.99 | 4.47 | 74.6 |
| 10 | 12.49 | 11.14 | 1.35 | 12.1 | 9.78 | 5.55 | 4.23 | 76.2 | 5.22 | 5.59 | — .37 | 6.6 |
| 11 | 9.66 | 7.11 | 2.55 | 35.9 | 5.37 | 2.87 | 2.50 | 87.1 | 5.17 | 4.24 | .93 | 21.9 |
| 12 | 2.88 | 4.49 | —1.61 | 35.9 | 2.06 | .51 | 1.55 | a | 3.78 | 3.98 | — .20 | 5.0 |
| 13 | 11.81 | 8.56 | 3.25 | 38.0 | 5.13 | 3.80 | 1.33 | 35.0 | 6.30 | 4.76 | 1.54 | 32.4 |
| 14 | 10.31 | 7.78 | 2.53 | 32.5 | 3.72 | 2.60 | 1.12 | 43.1 | 6.88 | 5.18 | 1.70 | 32.8 |
| 15 | 5.93 | 3.10 | 2.83 | 91.3 | 1.15 | 1.38 | — .23 | 16.7 | 4.11 | 1.72 | 2.39 | a |
| 16 | 5.88 | 6.69 | — .81 | 12.1 | 2.17 | 1.14 | 1.03 | 90.4 | 3.57 | 5.55 | —1.98 | 35.7 |
| 17 | 8.01 | 8.68 | — .67 | 7.7 | 3.13 | 2.15 | .98 | 45.6 | 4.45 | 6.53 | —2.08 | 31.9 |
| 18 | 3.25 | 4.31 | —1.06 | 24.6 | 1.89 | .80 | 1.09 | a | 2.06 | 3.52 | —1.46 | 41.5 |
| 19 | 2.16 | 2.49 | — .33 | 13.3 | .08 | .23 | — .15 | 65.2 | 1.95 | 2.26 | — .31 | 13.7 |
| 20 | 5.42 | 3.66 | 1.76 | 48.1 | .57 | .23 | .34 | a | 4.66 | 3.44 | 1.22 | 35.5 |
| 21 | 4.10 | 3.67 | .43 | 11.7 | 1.08 | .69 | .39 | 56.5 | 3.60 | 2.98 | .62 | 20.8 |
| 22 | 1.56 | 1.29 | .27 | 20.9 | 1.30 | .16 | 1.14 | a | 1.99 | 1.13 | .86 | 76.1 |
| 23 | 8.81 | 11.69 | —2.88 | 24.6 | 1.26 | .38 | .88 | a | 11.52 | 11.31 | .21 | 1.9 |
| 24 | 8.07 | 6.11 | 1.96 | 32.1 | .93 | .52 | .41 | 78.8 | 6.96 | 5.59 | 1.37 | 24.5 |
| 25 | 9.72 | 12.65 | —2.93 | 23.2 | 1.52 | 1.04 | .48 | 46.2 | 9.42 | 11.61 | —2.19 | 18.9 |
| 26 | 8.33 | 4.47 | 3.86 | 86.4 | 1.70 | .52 | 1.18 | a | 7.10 | 3.95 | 3.15 | 79.7 |
| 27 | 5.88 | 4.70 | 1.18 | 25.1 | .46 | .21 | .25 | a | 5.32 | 4.49 | .83 | 18.5 |
| 28 | 10.79 | 14.74 | —3.95 | 26.8 | .85 | .89 | — .04 | 4.5 | 9.99 | 13.85 | —3.86 | 27.9 |
| 29 | 13.75 | 11.39 | 2.36 | 20.7 | 1.03 | 1.42 | — .39 | 27.5 | 11.68 | 9.96 | 1.72 | 17.3 |
| 30 | 3.48 | 3.52 | — .04 | 1.1 | .21 | .20 | .01 | 5.0 | 3.49 | 3.33 | .16 | 4.8 |
| 31 | 4.35 | 2.43 | 1.92 | 79.0 | 1.42 | .90 | .52 | 57.8 | 3.24 | 1.53 | 1.71 | a |
| 32 | 5.41 | 10.18 | —4.77 | 46.9 | 2.66 | 7.55 | —4.89 | 64.8 | 3.48 | 2.62 | .86 | 32.8 |
| 33 | 4.13 | 3.47 | .66 | 19.0 | — .02 | .00 | — .02 | a | 2.93 | 3.47 | — .54 | 15.6 |
| 34 | 2.15 | 1.46 | .69 | 47.3 | — .68 | .02 | — .70 | a | 1.52 | 1.44 | .08 | 5.6 |
| 35 | 6.66 | 7.51 | — .85 | 11.3 | — .21 | .00 | — .21 | a | 5.25 | 7.50 | —2.25 | 30.0 |
| 36 | 7.61 | 9.21 | —1.60 | 17.4 | — .07 | .28 | — .35 | a | 6.23 | 8.93 | —2.70 | 30.2 |
| 37 | 8.98 | 10.29 | —1.31 | 12.7 | .14 | .22 | — .08 | 36.4 | 7.36 | 10.07 | —2.71 | 26.9 |
| 38 | 8.25 | 4.19 | 4.06 | 96.9 | .23 | .61 | — .38 | 62.3 | 8.40 | 3.58 | 4.82 | a |
| 39 | 12.05 | 14.34 | —2.29 | 16.0 | 1.42 | 1.64 | — .22 | 13.4 | 12.32 | 12.71 | — .39 | 3.1 |
| 40 | 9.83 | 13.32 | —3.49 | 26.2 | .57 | .59 | — .02 | 3.4 | 11.11 | 12.72 | —1.61 | 12.7 |
| 41 | — .43 | 1.13 | —1.56 | a | — .78 | .11 | — .89 | a | — .75 | 1.03 | —1.78 | a |
| 42 | 7.46 | 7.59 | — .13 | 1.7 | .61 | .52 | .09 | 17.3 | 6.08 | 7.07 | — .99 | 14.0 |
| 43 | 3.83 | 3.59 | .24 | 6.7 | — .75 | .07 | — .82 | a | 3.73 | 3.52 | .21 | 6.0 |
| 44 | 5.87 | 6.79 | — .92 | 13.5 | — .41 | .02 | — .43 | a | 4.37 | 6.77 | —2.40 | 35.5 |
| 45 | 6.26 | 5.75 | .51 | 8.9 | — .19 | .49 | — .68 | a | 4.56 | 5.26 | — .70 | 13.3 |
| 46 | 10.57 | 8.89 | 1.68 | 18.9 | 1.17 | 2.02 | — .85 | 42.1 | 8.20 | 6.87 | 1.33 | 19.4 |
| 47 | 7.57 | 8.34 | — .77 | 9.2 | — .27 | .76 | —1.03 | a | 7.08 | 7.59 | — .51 | 6.7 |

[a] Relative error is greater than 100 percent or undefined.

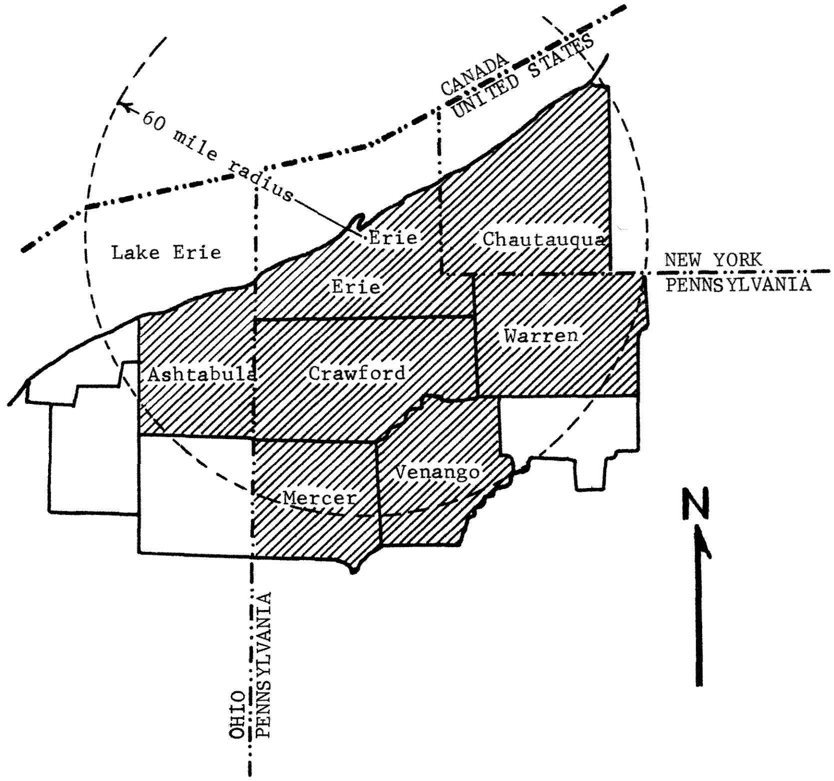


Figure 2.—Erie subregion.

## A CHECK ON PERFORMANCE

Does it work? This is the real measure of any predictive model's worth. To examine performance, we applied the equations for estimating local timber removals to the 47 sample subregions used in this study. Removals estimated by the equations were compared with recorded estimates (table 1). Absolute differences between recorded and equation estimates range from 0.01 to 6.02 cubic feet. Relative errors in the equation estimates, i.e., the absolute differences between equation estimates and recorded estimates expressed as a percentage of recorded estimates, range from 1.1 to more than 100 percent. The relative error in equation estimates is less for sample subregions in which recorded amounts of timber removed are large. This implies that the models should perform best where the amount of timber removed is known to be large. In some cases, equation estimates of removal were negative, but this occurs only where recorded removals are zero or negligible.

The above comparisons are by no means an exhaustive test of model performance, but they do indicate the kind of results that users of the equations might expect.

## AN APPLICATION

Now that we have the models how do we use them? Suppose we wanted to estimate the current annual removal of softwood, hardwood and all timber within a 60 mile radius of Erie, Pennsylvania (figure 2). This area is approximately the same as that included in Chautauqua county, New York; Ashtabula county, Ohio; and Erie, Crawford, Mercer, Warren, and Venango counties, Pennsylvania.

The first step is to determine values for the five regressor variables that are applicable to this problem. These values can be calculated very simply, using county data from the most recent Forest Survey and Census publications. For example, the value of $X_2$ (softwood growing stock inventory in thousands of cubic feet per acre of total land) is determined as follows:

$$X_2 = \frac{\text{million cubic feet of softwood growing stock}}{\text{thousand acres of total land}}$$

For our seven-county subregion:

| *County* | *Softwood growing stock (million cubic feet)* | *Total land area (thousand acres)* |
|---|---|---|
| Chautauqua, N. Y. | 42.9[1] | 691.6[1] |
| Ashtabula, Ohio | .6[2] | 448.1[2] |
| Crawford, Pa. | 17.3[3] | 650.2[3] |
| Erie, Pa. | 14.5[3] | 519.7[3] |
| Mercer, Pa. | 8.0[3] | 435.8[3] |
| Venango, Pa. | 39.5[3] | 432.0[3] |
| Warren, Pa. | 55.9[3] | 582.4[3] |
| Subregion total | 178.7 | 3,759.8 |

$$X_2 = \frac{178.7}{3{,}759.8} = 0.048$$

Similarly, values for the other four regressor variables are:

$X_3$ =Hardwood growing stock inventory in thousands of cubic feet per acre of total land[1, 2, 3]
=0.488

$X_4$ =Softwood sawtimber volume inventory in thousands of board feet per acre of total land[1, 2, 3]
=0.090

$X_8$ =Number of lumber and wood-product firms (SIC 24) per thousand acres of total land[1, 2, 3, 4]
=0.052

$X_{15}$=Per capita income in thousands of dollars[5]
=2.777

Now, we substitute these values into regression equations (2) and (3) to solve for $Y_S$ and $Y_H$, respectively:

$Y_S$=The current annual removal of softwood timber
$=-1.29+17.02X_2+28.99X_8$
$=-1.29+17.02(0.048)+28.99(0.052)$
=1.04 cubic feet per acre of total land, and

$Y_H$=The current annual removal of hardwood timber
$=9.55+5.53X_3+63.01X_8$
$-2.99X_{15}-1.93X_4$
$=9.55+5.53(0.488)+63.01(0.052)$
$-2.99(2.777)-1.93(0.090)$
=7.06 cubic feet per acre of total land.

Since $Y_S$ and $Y_H$ are per-acre values, we multiply them by the total acreage of land in the subregion to get estimates of the total annual removals of softwood and hardwood timber in the subregion. Thus: the estimated total annual removal of softwood timber=1.04 (3,759,800) =3.9 million cubic feet, and the estimated total annual removal of hardwood timber=7.06 (3,759,800) =26.5 million cubic feet.

These estimates of the softwood and hardwood timber removals can be added to get an estimate of the annual removal of all timber in the subregion: the estimated total annual removal of all timber=3.9+26.5=30.4 million cubic feet.

An alternative method of estimating the annual removal of all timber is to solve regression equation (1) for $Y_A$ and then multiply $Y_A$ by the total acreage of land in the subregion:

$Y_A$=the current annual removal of all timber
$=15.13+13.70X_2+116.49X_8$
$-4.78X_{15}-2.25X_4$
$=15.13+13.70(0.048)+116.49(0.052)$
$-4.78(2.777)-2.25(0.090)$
=8.38 cubic feet per acre of total land

[1] Source: Ferguson, Roland H. and Carl E. Mayer. 1970. THE TIMBER RESOURCES OF NEW YORK. USDA Forest Serv. Res. Pap. NE-20, 193 p.

[2] Source: Kingsley, Neal P. and Carl E. Mayer. 1970. THE TIMBER RESOURCES OF OHIO. USDA Forest Serv. Res. Pap. NE-19, 137 p.

[3] Source: Ferguson, Roland H. 1968. THE TIMBER RESOURCES OF PENNSYLVANIA. USDA Forest Serv. Res. Pap. NE-8, 147 p.

[4] Source: U.S. Bureau of the Census. 1971. CENSUS OF MANUFACTURES: 1967, Volume III: Area Statistics. U.S. Government Printing Office, Washington, D. C.

[5] Source: U.S. Bureau of the Census. 1972. CENSUS OF POPULATION: 1970: General Social and Economic Characteristics. PC(1)-C34 P. Y., PC(1)-C37 Ohio, and PC(1)-C40 Pa. U.S. Government Printing Office, Washington, D. C.

Then the estimated total annual removal of all timber in the subregion=8.38 (3,759,800) =31.5 million cubic feet. The difference between the two alternative estimates of the total annual removal of all timber=31.5–30.4 =1.1 million cubic feet. In this case the difference is less than 4 percent of either estimate. Results will vary from case to case.

## SUMMARY

Resource analysts have long needed a practical method of estimating current removals of timber for local areas in the Northeast. In response to this need, we have developed regression equations for estimating subregional removals of softwood, hardwood, and all timber as a function of selected characteristics of socioeconomic structure. Five key regressor variables are included in the equations. These variables reflect conditions of timber availability, market activity and demand, and land-use competition that determine the amount of local timber removed.

The regression models can be used to estimate timber removals for any subregion of the Northeast. But they are likely to perform best for areas of a million acres or more where the quantity of timber removed is known to be large.

---