

Pre-harvest Pecan Yield Estimation

Glenn C. Wright¹, J. Benton Storey², Marvin K. Harris³, and Peter T. Sprinz⁴

Texas A&M University, College Station, TX 77843

Additional index words. *Carya illinoensis*, computer applications, sampling, statistical analysis

Abstract. Nut count (NC), trunk circumference (TC), competition factor (CF), days from budbreak (DAY), and high or low crop year (YR) data were collected on 40 trees at three sites across Texas in 1985 and 1986, to create a model that would predict pecan [*Carya illinoensis* (Wangenh.) C. Koch] yield. The model developed predicted the natural logarithm of the total nuts on the tree [ln(NUTS)]: $\ln(\text{NUTS}) = 2.112 + [0.634 \times \ln(\text{NC})] + (0.00119 \times \text{TC}) - (0.0701 \times \ln(\text{CF})) + (0.00639 \times \text{DAY}) + (0.728 \times \text{YR})$. The equation accounts for 87% of the variation in yield. The model is not sufficiently accurate to predict individual tree yields well, but additional data show an ability to accurately predict average tree yields.

Yield estimates for pecan are often difficult because of the variation in tree size, the variation in nut population throughout the year, and the inability to distinguish between small green nuts and the surrounding leaves in the upper portions of the trees. Pre-harvest yield data would be useful for making integrated pest management decisions, where estimates of final harvest data could be used to determine if a cultural practice was economically feasible, and for establishing appropriate prices.

Limited research has been conducted on the components of pecan yield (Malstrom and Sparks, 1973; Worley et al., 1972) and we found no research on estimating pecan yield using regression models. Some yield estimates have been made for deciduous fruit crops, chiefly sweet cherries (*Prunus avium* L.) and apples (*Malus domestica* Borkh.). Many of these studies have indicated the tree characteristics are associated with fruit yield. Forshey and Elfving (1979) suggested branch sampling might be appropriate. Lozano-Gonzalez (1982) found pecan branch circumference at an attachment point gave the best estimate of the number of nuts on that branch. Trunk size has also been used to estimate yield. Lutz (1938) noted a close relationship between tree size (estimated by measuring trunk cross-sectional area) and yield. Worley et al. (1972) reported that pecan trunk circumference was positively associated with yield high crop years, while

there was no significant correlation in low crop years.

Yield may also be affected by competition among the trees. Hinrichs (1961) found that reducing dense stands increased production and concluded that yields were inversely related to cross-sectional trunk area per acre as optimum stand densities were exceeded. Foresters have developed a method of quantifying competition into values known as point density measurements that are attempts to incorporate varying degrees of competition experienced by individual trees into the procedure of predicting individual tree growth (Clutter et al., 1983). These values are based on the diameter of the subject trees and on the diameters of and distances to competing trees.

The objective of this experiment was to collect data on nut population, trunk circumference, and competition between trees, and from these values develop an accurate and feasible model to predict, throughout the season, ultimate harvestable yield.

Nineteen trees of 14 cultivars were sampled in 1985. All trees were located on the Texas A&M Univ. plantation, 16 km southwest of College Station in the Brazes River flood plain. Fertilization regime included a single application of 102 kg N/ha as granular NH_4NO_3 ; two applications of 2.63 kg N/ha and 0.66 kg Zn/ha as a liquid fertilizer containing $\text{Zn}(\text{NO}_3)_2$ and a urea plus ammonium nitrate solution; and a single application of 4.7 kg Zn/ha as ZnSO_4 . Pest control included the use of cyano(3-phenoxyphenyl)methyl-4-chloroalpa-(1-methyl-ethyl)benzeneacetate (fenvalerate) and S[(6-chloro-2-oxo-3 (2H)-benzoxazolyl)methyl]0,0-diethylphosphorodithioate (phosalone) for insect control; methyl[1-(butylamino)carbonyl]-1H-benzimidazol-2-yl]-carbonate (benomyl) and triphenyltin hydroxide (fentin hydroxide) for disease control, and 6-chloro-*N,N'*-diethyl-1,3,5-triazine-2,4-diamine (simazine), 1,1'-dimethyl-4,4'-bipyridinium salts (paraquat), and N-(phosphonomethyl)glycine (glyphosate) for control of weeds, at appropriate times throughout the season.

Trees were selected for inclusion in the

experiment based on visual appearance of health. The total population of branches on each tree was stratified, based on normal vigor (15 to 45 cm of terminal growth), location in the upper half of the canopy, as recommended by Malstrom and Sparks (1973) and Lozano-Gonzalez (1982) for large trees, location in the center of the canopy for small trees, and a diameter of 3.5 to 10 cm at the attachment point (Lozano-Gonzalez, 1982). [Snedecor (1946) states that stratification is "...subdividing the aggregate into relatively homogeneous strata. . . a device for utilizing any knowledge one may have of the outcome in order to exclude from the estimate of error, that variation which he can anticipate"]. Branches not meeting these criteria were not considered. Six branches per tree were sampled at random from the remaining population.

The circumference of each branch selected (LC) was measured at its attachment point to a scaffold limb. All nuts distal from the point of attachment were counted. NC was the average number of nuts counted on each of the six branches sampled. Following nut counting, trunk circumference (TC) was measured on each tree at ≈ 137 cm above the ground. Trees with branches at or below this point were measured just below those branches (Dewars and Driessen, 1983).

Trunk diameters of all trees within a radius of 22.9 m from the subject tree, as well as the actual distance from the subject tree to those competing trees, were measured. These data were used in calculating the competition factor (CF), developed by Heygi (1974).

$$\text{DCI}_j = \sum_{i=1}^n [(D_i/D_j) \times (1/\text{DIS}_{ij})] \quad [1]$$

where: DCI_j is the competition index of the j th subject tree; D_i is the diameter of the i th competitor; D_j is the diameter of the j th subject tree; DIS_{ij} is the distance between the j th subject tree and its competitor; and n is the number of competitors of the j th subject tree within a circle of 22.9 m. CF was calculated for all subject trees. When a tree was located in an area where there were < 7.65 m² of cross-sectional trunk area per hectare, competition was considered to not affect yield (Hinrichs, 1961) and the index was set at 0.000001.

Trees were harvested in Oct. and Nov. 1985. All nuts harvested from the 19 trees were bagged and processed. Foreign material, such as sticks and leaves, as well as broken pecans were removed during processing, leaving all whole pecans, regardless of quality, in the bag. The nuts were allowed to dry for 30 days to reach a constant weight (Storey, 1958) and weighed. A 2.25-kg sample was then removed from each bag. Twenty nuts were selected at random from the sample, and the average weight per nut was determined. The total weight was then divided by the average weight per nut and an estimate of the total number of nuts per tree (NUTS) was determined.

Received for publication 9 Jan. 1989. This research was supported in part by USDA Agreement no. 85-CSRS-2-2568 from CSRS. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact.

¹Graduate student, Dept. of Horticultural Sciences.

²Professor of Horticulture, Dept. of Horticultural Sciences.

³Professor of Entomology, Dept. of Entomology.

⁴Assistant Professor of Biometrics, Dept. of Forest Science. Present address: Eastman Kodak Inc., Rochester, NY 11177.

Table 1. Parameter estimates for the combined 1985-1986 Model for predicting pecan yield: $\ln(\text{NUTS}) = f[\ln(\text{NC}), \ln(\text{CF}), \text{DAY}, \text{YR}]$ ($R^2 = 0.869$).

Variable estimate	Parameter	SE	Significance level ¹
Intercept	2.112	0.537	0.0002
lnNC	0.634	0.105	0.0001
TC	0.0012	2.3×10^{-4}	0.0001
lnCF	-0.0701	0.0167	0.0001
DAY	0.0063	0.0017	0.0004
YR	0.729	0.249	0.0047

¹NC = nutcount; TC = trunk circumference; CF = competition factor; DAY = days from budbreak YR = high or low crop year.

²Based on $t_{0.025,76} = 1.990$.

Table 2. R^2 and MSE statistics for all models for predicting pecan yield.

Model	R^2	M	S	E
Sept. 1985 $\ln(\text{NUTS}) = f(\text{NC}, \text{LC}, \ln\text{TC}, \text{CF})$	0.8570	0.077		
May 1985 $\ln(\text{NUTS}) = f(\ln\text{NC}, \text{LCTC}, \ln\text{CF})$	0.5850	0.249		
July 1986 $\ln(\text{NUTS}) = f(\ln\text{NC}, \text{LCTC}, \text{CF})$	0.8640	0.081		
Sept. 1986 $\ln(\text{NUTS}) = f(\ln\text{NC}, \text{LCTC}, \ln\text{CF})$	0.8510	0.089		
Combined Model $\ln(\text{NUTS}) = f(\ln\text{NC}, \text{TC}, \ln\text{CF}, \text{DAY}, \text{YR})$	0.8690	0.187		

Table 3. Estimated and actual pecan yield of trees sampled in 1987.

Actual yield (kg)	Estimated yield (kg)	Difference (d) (kg)
12.734	12.739	-0.024
6.630	9.004	-2.374
2.268	6.428	-4.161
19.428	16.212	3.216
19.591	16.414	3.177
9.347	10.398	-1.052
14.235	10.637	3.599
10.000	11.412	-1.412
11.428	12.696	-1.268
\bar{X} 11.740	11.773	2.604 ²

²Means are not significantly different from each other at $P = 0.05$, based on $t_{0.05,8,8} = 1.860$.
³ $(d_i^2/n)^{1/2}$.

Data were collected in May 1986 at the university orchard and at two privately owned orchards near Seguin and Uvalde, Texas. Twenty-one trees comprising four cultivars and native trees were used in 1986. Cultural practices at the university and Seguin orchards were the same as those previously mentioned. Practices in Uvalde included a single application of 112 kg N/ha as $(\text{NH}_4)_2\text{SO}_4$ and five applications of 2.05 kg N/ha and 0.51 kg Zn/ha as a liquid fertilizer containing $\text{Zn}(\text{NO}_3)_2$ and a urea plus ammonium nitrate solution. Pest control in Uvalde included the use of phosalone for insect control, benomyl for control of stem and blight, and simazine, paraquat, and glyphosate for weed control. Supplemental irrigation in Uvalde was applied to deliver the equivalent of 70% of the daily Class A pan evaporation per day. Branches were selected,

measured, and the nut population counted in May. LC and NC were recorded again, on the same branches, in July and September. TC was measured and CF calculated in September. Pecans were hand- or machine-harvested, processed, and weighed in the same manner as in 1985.

The models constructed from the data for each sampling period are denoted as September 1985, May 1986, July 1986, and September 1986. The natural logarithm of the nuts per tree [$\ln(\text{NUTS})$] was the dependent variable. NUTS was chosen rather than total yield because of the variability of nut size between cultivars. NUTS is also a better indicator of the total host for insects or disease than its total yield. Untransformed and transformed NC, LC, TC, and CF were the independent variables in each case.

Limitations of these four models were quickly noticed. Each model was only usable during that month in which the original data were collected, and only during the appropriate high or low crop year. Therefore, data from both 1985 and 1986 were combined into one data set for the construction of a fifth model, sensitive to the time of the year, known as the Combined Model. Appropriate transformed and untransformed variables (NC, TC, and CF) were also included in this model; however, limb circumference (LC) was not included because t tests for this variable in the interim models indicated no significant contribution.

Another variable, days from budbreak (DAY), was included in the Combined Model, so that the resulting model could be used throughout the growing season. Budbreak occurred on 15 Mar. at Uvalde, 25 Mar. at Seguin, and 1 Apr. at College Station. DAY was calculated using these dates as starting points.

Pecans are usually alternate bearing, so high crop years usually follow low crop years if natural disasters do not intervene. High crops are obvious when casual observation indicates > 90% of the terminals have nuts. Casual observation denotes practically no crop in low crop years. High or low crop year was included in the model as the variable YR, having a value of 1 or 0, respectively.

Statistics generated for evaluating the models included the coefficient of determination (R^2), and mean square error (MSE).

Additional data were collected in 1987 to test the model at a privately owned orchard in Comanche County, Texas. NC, TC, and CF were determined for 'Cherokee' trees. DAY was 180 and YR was 1. Harvest began 24 days following sampling and nuts were treated in the manner described above. Determination of significant differences between mean actual and mean model-estimated yields was by a t test for paired data.

The best model for the combined 1985-1986 data was:

$$\ln(\text{NUTS}) = 2.112 + [0.634 \times \ln(\text{NC})] + (0.00119 \times \text{TC}) - [0.0701 \times \ln(\text{CF})] + (0.00639 \times \text{DAY}) + (0.728 \times \text{YR}) \quad [2]$$

The R^2 of this Combined Model was 0.869 and the MSE was 0.187. Parameter estimates and SE for the Combined Model are listed in Table 1.

MSE and R^2 for the four interim models and the Combined Model (Eq. [2]) are listed in Table 2. The May 1986 model had a lower R^2 , probably due to the inability of the independent variables, especially NC, to account for the variability introduced by June nut drop, on the final yield. The Combined Model had a higher R^2 than any of the other models, probably due to improved data fit over a wider range of data. The R^2 for the September 1985 model was similar to those of the July and September 1986 models and the Combined Model, indicating high goodness of fit. Model prediction and data fit improved after the June nut drop period, as shown in the improved July and September 1986 models. The similar statistics of the July 1986 and September 1986 models indicated little difference in their abilities to fit the data and predict nut populations.

Limited data collected in 1987 were used for testing purposes (Table 3). Model-estimated yield varied from actual yield on the nine trees by as much as 4.1 kg. Mean estimated yield was not statistically different from mean actual yield at $P = 0.05$. The Combined Model is apparently more accurate in predicting general trends and mean yields than specific yields for individual trees because of the high variability of those trees. Sampling regime for orchards should, therefore, consist of trees from a fairly homogeneous population. Additionally, the accuracy of the Combined Model will depend on the number of trees sampled and the number of branches sampled per tree. The sample size can be calculated using a formula presented by Southwood (1971). Generally, a sample size of 10 to 24 trees should be sufficient. Additionally, a proper sampling regime should include at least five branches per tree (Wright, 1987).

Satisfactory predictions of expected yield were obtained with some models. The May 1986 model was the least accurate, while the July 1986, September 1986, and September 1985 models were more accurate for estimation of final yield. The Combined Model was built with all the data, including days from budbreak variable and a code for high and low crop year, and was the most useful model, since it eliminated the need for a separate model for each sampling time and year.

Limited additional data indicate that the Combined Model performs well (Table 3). It represents the first step toward a reliable technique to estimate yield for many decision-making purposes throughout the growing season.

Pecan yields are highly variable from location to location and year to year with harvest densities ranging from 0 to several thousand kilograms per hectare. Verification of this model would require data from many sites, cultivars, and conditions, over several seasons. New variables, such as one for cultivar, might be added to refine the model. It is important that growers and pecan scien-

tists use this model, establish its validity for their specific locations, and tailor it to their own use.

Literature Cited

- Clutter, J. L., J.C. Forston, J.V. Pineaar, G.H. Bristar, and R.L. Bailey. 1983. Timber management: A quantitative approach. Wiley, New York.
- Dewars, R.S. and A.D. Driessen. 1983. Evaluation of Texas shade trees. Texas Agr. Ext. Serv. Fact Sheet L-1683.
- Forshey, C.G. and D.C. Ehling. 1979. Branch samples for yield and fruit size comparisons in apple. HortScience. 14:143-144.
- Heygi, F. 1974. A simulation model for managed jackpine stands. Growth models for trees and stand simulation. Royal College For. Res. Notes no. 30. Stockholm.
- Hinrichs, H.A. 1961. The relationship of native pecan tree spacing to yield. Okla. Agr. Expt. Sta. Bul. B-574.
- Lozano-Gonzalez, R. 1982. A three-dimensional characterization of a bearing pecan tree. MS Thesis. Texas A&M Univ., College Station.
- Lutz, H. 1938. The 'effect of size of pecan trees on their subsequent growth and yield. Proc. Amer. Soc. Hort. Sci. 36:335-338.
- Malstrom, H.L. and D. Sparks. 1973. Analysis of yield components in mature trees in 'Schley' pecan, *Carya illinoensis* (Koch). J. Amer. Soc. Hort. Sci. 98:496-500.
- Snedecor, G.W. 1946. Statistical methods. 4th ed. Iowa State Univ. Press, Ames.
- Southwood, T.R.E. 1971. Ecological methods. 3rd ed. Chapman and Hall, London.
- Storey, J.B. 1958. Moisture loss from pecans during storage. Proc. Texas Pecan Growers Assn. 37:44-47.
- Worley, R. E., S.A. Harmon, and R.L. Carter. 1972. Correlation among growth, yield, and nutritional factors for pecan (*Carya illinoensis* W. cv. 'Stuart'): Correlations with yield, quality, and terminal shoot growth. J. Amer. Soc. Hort. Sci. 97:511-514.
- Wright, G.C. 1987. Pre-harvest pecan yield estimation in Texas. MS Thesis. Texas A&M Univ., College Station.

HORTSCIENCE 25(6):700-704. 1990.

Interregional Analysis of Fresh Dry Onion Production: An Examination of Texas in the Spring and Summer

Stephen Fuller¹, H.L. Goodwin², and Carl Shafer¹

Department of Agricultural Economics, Texas A&M University, College Station, TX 77843

Grant Vest³

Department of Horticultural Science, Texas A&M University, College Station, TX 77843

John Schmitz⁴

Department of Agricultural Economics, Texas A&M University, College Station, TX 77843

Additional index words. economic analysis, *Allium cepa*, marketing

Abstract. A national trade model of the fresh onion (*Allium cepa* L.) market is used to identify the capability of Texas to compete with other production areas in dry onion sales throughout the United States. The analysis shows Texas and western U.S. producers to be major competitors in the spring and summer seasons. While Texas producers' proximity to major eastern markets yields a cost advantage, their market share is vulnerable to cost reductions by western U.S. shippers. Thus, both cost-reducing innovations (production) and quality improvements are important in maintaining the competitiveness of Texas producers.

The declining market share of the spring and summer fresh dry onion market supplied by onion producers in the subtropical portion of the Rio Grande Valley (Texas) and South Plains (Fig. 1) has created uncertainty about the producers' ability to remain competitive. A national interregional trade model was used to identify the ability of the Rio Grande Valley of Texas and the South Plains to compete for dry onion sales in various geographic regions of the United States.

Received for publication 1 Aug. 1989. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact.

¹Professor.

²Associate Professor.

³Professor and Department Head.

⁴Graduate Research Assistant.

Dry onion production in the United States is identified as 1) spring, 2) summer non-storage, or 3) summer storage crop. The spring crop comprises 15% to 18% of U.S. onion production and is produced primarily in California, Arizona, and the lower Rio Grande Valley of Texas. Spring onions are the first-harvested of the calendar year and typically move directly to the fresh market. Summer non-storage onions, ≈10% of national production, are produced primarily in the upper Rio Grande Valley (Texas and New Mexico). Summer storage onions, 70% to 75% of U.S. onion production, are harvested in northern states from August through October, with shipments continuing through early spring (April). Storage onions do not compete with new onion production until spring harvest commences in March; however, imports from Mexico offer competition during

the winter and early spring.

Per-capita consumption of fresh onions has increased 33% during the past decade, from 6.1 kg in the mid-1970s to nearly 8.2 kg per capita by 1987. The increase is attributed in part to health-conscious consumers increasing consumption of fresh vegetables and away-from-home consumption, which often includes both fresh and processed onion products. Per-capita consumption of onions has also benefited from increased consumption of ethnic foods, particularly Mexican.

While the U.S. production and consumption of dry onions was increasing during 1977-1987, shipments from Texas decreased modestly. Thus, market share for Texas producers declined. Monthly market-share trends calculated for the April-August window showed the lower Rio Grande Valley's national share in April remained relatively constant at ≈70%. In contrast, market share in May and June declined ≈40% over the 1977-1987 period, with the current monthly shares averaging ≈40% and 15%, respectively. Similar declines in onion market share were observed for the summer months, a period when shipments originate in the South Plains. Much of Texas producers' declining market share was claimed by California producers. Competing regions other than California showed no definite trend in market share. Whether onion producers in Texas found alternative enterprises that were more profitable or experienced resource constraints that fixed upper bounds on onion production was not determined.

Framework for analysis

Efficient resource use implies a production and geographic distribution that satisfies market requirements at the lowest possible cost (Bressler and King, 1970). King (1969) suggested that a transportation model, in combination with production and processing costs, offers insight into the cost-competitiveness of producing regions. Several studies have used this framework successfully (Bates and Schmitz, 1969; Ben-David and Forker, 1970; O'Rourke, 1977).

Our study uses an interregional transportation model to address the competitiveness issue. The model uses each region's monthly