Randomized branch sampling to estimate fruit production in Pecan trees cv. ‘Barton’

Filemom Manoel Mokochinski¹* Luciano Farinha Watzlawick¹ Renato Vasconcelos Botelho¹ Vanessa Moreira¹

¹Departamento de Agronomia, Universidade Estadual do Centro-Oeste (UNICENTRO), 85040-080, Guarapuava, PR, Brasil. E-mail: filemom_mom@hotmail.com. *Corresponding author.

INTRODUCTION

In recent decades, there has been an increased demand for healthier foods that have the nutritional elements recommended for health maintenance, such as nuts. The pecan tree (Carya illinoinsensis (Wangenh.) K. Koch) provides fruits that are considered an important raw material for agro-industry. Its almonds can be used in the food industry and for the production of oil. Shells can be used to obtain tea and fertilizers, and its wood can be used for furniture. The species also enables a consortium with other cultures and the integration of livestock.

Despite these benefits, the pecan crop has not been encouraged, since little research has been conducted on its implementation and management (DUARTE & ORTIZ, 2001). Studies that seek to assist producers in pecan orchard planning should be prioritized.

Methods that quantify the productive potential of orchards, based on reliable sampling techniques, are practically nonexistent. Sampling methods may provide an alternative to the census by estimating the attributes of interest in individual trees. This may prove to be an extremely useful tool, allowing producers to be informed on whether the chosen species will have good productive output.

JESSEN (1955) developed the technique of randomized branch sampling (RBS) to estimate the attributes of interest in individual trees. The technique is based on the selection of branches in...
the tree and randomly assigning probabilities to that selection of branches. In the uniform probability technique (UP), all sampled branches have the same probability, independent of its diameter, whereas in the probability proportional to the diameter technique (PPD), the branch diameter is taken into account in the selection of the final branch.

In the RBS technique, a single tree is regarded as the population, and the target being measured is the total amount of the quantified attributes in the sampled branches. Branch sampling has been used to assess the amount of fruits, seeds, insect attack and for the quantification of volume, biomass and carbon stocks (BORGES, 2009; RYALL et al., 2011; BARBEIRO, 2012; LÔBO, 2012).

So, the goal of this study was to determine the efficiency of probability sampling of branches to estimate pecan production at three different ages by comparing techniques of uniform probability and probability proportional to the diameter, with the overall aim to support farmers in the assessments of their orchards.

MATERIALS AND METHODS

The study area is located in the county of Cacheoeira do Sul - RS, Jacuí Valley region, with the geographical coordinates of 30°20'16" and 30°20'25" South and 52°00'23" and 52°03'27" West, at an average altitude of 60 meters. According to the climatic classification of KÖPPEN (1948), the region has a Subtropical Climate (CFA), whose main characteristics include winters with the occurrence of severe and frequent frosts, an average annual temperature of 19°C and an annual rainfall of approximately 1500mm year with no dry season.

Pecan plantations were of three different ages (5, 7 and 10 years) and had been cultivated under the same management practices since their implementation, with a spacing of 7x7 meters. To select the trees, three individuals from each age were randomly selected. Each sampled tree was regarded as a population, which was evaluated using the techniques uniform probability (UP) and probability proportional to the diameter (PPD). Fruits from the trees were the estimated attributes.

Branch sampling technique uses the natural offshoot from the inside of the crown to compose its sample, requiring prior establishment of branches, nodes, segments of branches and paths. Branches form an entire system that has its source in a single bud, lateral or terminal, and comprises a subpopulation within the tree. The node is the point on the tree where a branch splits into two or more (bifurcation). Segments are the part of a branch between two consecutive nodes, and the path is a series of segments between the basal segment and the terminal bud. Number of possible paths in a tree is equal to the number of terminal buds.

The first segment of the path extends from the base of the stem, defined as the first node, to the second node. Most of the time, the first node of the path leads to one branch (in this case, the stem), where the probability of selection will be equal to one \((q_1 = 1)\). In the second node, the probability of selection will be assigned to each branch, randomly choosing one. The choice for the second branch, with the probability of selection \(q_2\), defines the second segment of the path. This procedure is repeated until a lower branch or an end segment is selected at the final node, with probability \(q_{ri}\).

To assign probabilities proportional to the diameter for each branch, the following formula was used, adapted from GREGOIRE & VALENTINE (2007). The diameter of the branches was taken near the node and branches were measured with a digital caliper:

\[
q_{ri} = \frac{d_{ri}^2}{\sum_{i=1}^{n} d_{ri}^2} \quad i = 1, 2, ..., n
\]

where:
- \(q_{ri}\) = probability of selection assigned to the \(i\)-th branch from the \(r\)-th node;
- \(d_{ri}^2\) = squared diameter \((d^2)\) of the \(i\)-th branch from the \(r\)-th node;
- \(\sum_{i=1}^{n} d_{ri}^2\) = sum of the squared diameters of the \(i\)-th branch from the \(r\)-th node.

Thus, the branch is selected if \(u < q_{ri}\), where \(u\) is a random number between 0 and 1 that will determine which branches at a particular node will be chosen. For the selection of branches, \(u\) values with three decimal places were determined, and the branches were randomly assigned a number between 0 and 1 (e.g., \(u = 0.123\)). The draw of the three decimal places was done with 10 cubes, numbered from 0 to 9. When the raffled \(u\) value was less than \(q_{ri}\), the branch was chosen.

To assign a uniform probability (UP) for selecting branches at each node of the path, the following formula was used:

\[
q_{ri} = \frac{1}{n}
\]

where:
- \(n\) = number of branches from the \(r\)-th node.

To estimate the number of fruits on each tree, the number of fruits were counted in the segments.
Randomized branch sampling to estimate fruit production in Pecan trees cv. ‘Barton’.

of five paths on the selected terminal branch, and the calculation of unconditional probability of selection \((Q_{mi})\) was performed. These calculations allowed the estimation of the total amount of fruit on each tree through the sum of all sampled segments using the following formula:

\[
y_{Qm} = \frac{Y_{mi}}{Q_{mi}}
\]

where:
- \(y_{Qm}\) = estimate of the number of fruits of the tree given by the path \(m\);
- \(Y_{mi}\) = number of fruits from the branch \(i\) in the path \(m\);
- \(Q_{mi}\) = unconditional probability of the selection of the branch \(i\) in the path \(m\).

After obtaining an estimate of the number of fruits of \(m\) paths \((m \geq 2)\), these estimates were combined to estimate an average number of fruits per tree \((\bar{y}Q)\), which is given by:

\[
\bar{y}Q = \frac{1}{m} \sum_{j=1}^{m} y_{Qm}
\]

The joint probability of the selection of all segments \(S\) forming the \(m\)-th path is represented by \(Q_{mi}\). GREGOIRE & VALENTINE (2007) stated that, though the number of segments \((S)\) can vary between paths, if there are \(M\) possible paths with different terminal segments, the sum of the unconditional selection probabilities of all possible branches in a tree will equal one. Thus, we assumed that \(Q_{mi}\) is the probability of obtaining the estimate \((y_{Qm})\), which represents one of \(M\) possible estimates of fruits for the sampled tree.

For all sampled trees, the coefficient of variation \((cv)\) and the relative sampling error \((Er)\) were calculated:

\[
cv = \frac{S_y}{\bar{y}Q} \times 100
\]

\[
Er = \frac{S_y}{\bar{y}Q} \times 100
\]

where:
- \(cv\) = estimate of the coefficient of variation;
- \(\bar{y}Q\) = average number of fruits;
- \(S_y\) = estimate of the standard deviation.

\[
ER (%) = \left(\frac{cv_{y1}}{cv_{y2}}\right) \times 100
\]

where:
- \(ER\) (%) = relative efficiency

RESULTS AND DISCUSSION

Estimates of pecan fruit production using the PPD and UP techniques regarded the tree as a population of possible paths, which resulted in five production estimates for each technique. In all sampled paths, both techniques underestimated fruit production averages for trees (Table 1).

The number of fruits reported in the sampled terminal branches averaged 3.5 fruits per tree at 5 years, 18.6 fruits at 7 years and 19.2 fruits at 10 years. There was a 81.2% increase in the number of fruits as the age increased from 5 to 7 years and a 3.12% increase from 7 to 10 years. The small difference in the average of fruits on the sampled branches at 7 and 10 years demonstrates the stabilization trend in the number of fruits in the terminal branches as trees age.

The time required to perform the PPD was, on average, five-fold higher than the UP. The time difference for carrying out the PPD is related to the difficulty in measuring the diameters. According to WOLF (2009), the specific characteristics of growth and the arrangement of branches and crown affect the sampling time due to trees irregularity.

In some paths, few fruits were reported in the terminal branch, resulting in an underestimation of the real fruit production. Terminal branches without any fruits were also selected, which contributed to the overall underestimation. A similar situation was reported by BORGES (2009), who quantified fruit production of pequi and reported that branches with no fruit on fruiting trees resulted in the underestimation of mean production.

Trees at 10 years had the highest production, at approximately 1,559.30 units. Special attention was given to some of the terminal branches that presented the same number of fruits, but had different estimates because of the probability of selection \((Q_{mi})\). Coefficients of variation \((cv)\) for both sampling techniques showed oscillation in different paths. For trees at 7 years using the UP, and at 10 years using the PPD, the coefficient of variation decreased by increasing the sample size.
Analyzing real fruit production using the census, trees at 7 years produced approximately twelvefold more fruit than trees at 5 years, indicating that the beginning of pecan production takes 4 or 5 years, and at 7 years, production reaches commercial levels.

By analyzing only the sampling errors produced by the estimates from RBS, we can observe that the values were high. We noticed that the sampling error (%) decreased with an increase in the sample size. Regardless of the technique used (PPD or UP), the more paths sampled, the smaller the errors produced by the RBS (Figure 1). Similar results were reported by other authors, who suggested increasing the sample size to reduce the variance and errors (GOOD et al., 2001; CANCINO, 2005; BORGES, 2009).

The mean sampling error was 125.17% for the PPD and 111.04% for the UP. However, when the UP was used on trees at 7 and 10 years, assuming a sample size of 5 paths, the sampling errors were less than 10%. That is, in larger trees taking five paths, the sampling errors are smaller.

The largest errors were reported in a sample size of 10 and 1118.5 fruits. The mean sampling error was 125.17% for the PPD and 111.04% for the UP. However, when the UP was used on trees at 7 and 10 years, assuming a sample size of 5 paths, the sampling errors were less than 10%. That is, in larger trees taking five paths, the sampling errors are smaller.

### Table 1 - Mean variables of random sampling of branches performed on pecan trees.

<table>
<thead>
<tr>
<th>M</th>
<th>Fruits/branch</th>
<th>Qmi</th>
<th>țiQm</th>
<th>țiQ</th>
<th>Time (s)</th>
<th>cvy (%)</th>
<th>Er (%)</th>
<th>Real Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPD</td>
<td>5 years</td>
<td>1</td>
<td>1</td>
<td>0.0263</td>
<td>38</td>
<td>~</td>
<td>237.9</td>
<td>96.5 fruits</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>0.0211</td>
<td>94.5</td>
<td>66.2</td>
<td>294.9</td>
<td>61.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>7</td>
<td>0.0813</td>
<td>86</td>
<td>72.8</td>
<td>195.6</td>
<td>22.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>0.0707</td>
<td>56.5</td>
<td>68.7</td>
<td>247.8</td>
<td>15.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>2</td>
<td>0.0645</td>
<td>31</td>
<td>61.2</td>
<td>318.3</td>
<td>33.30</td>
</tr>
<tr>
<td>Mean</td>
<td>3.2</td>
<td>5</td>
<td>0.0509</td>
<td>78.5</td>
<td>~</td>
<td>287.1</td>
<td>33.21</td>
<td>167.97</td>
</tr>
<tr>
<td>UP</td>
<td>5</td>
<td>2</td>
<td>0.0740</td>
<td>27</td>
<td>52.7</td>
<td>49.5</td>
<td>40.11</td>
<td>360.24</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>0.0344</td>
<td>87</td>
<td>64.1</td>
<td>53</td>
<td>27.93</td>
<td>70.98</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>0.0457</td>
<td>87.5</td>
<td>70</td>
<td>112</td>
<td>28.50</td>
<td>45.29</td>
</tr>
<tr>
<td>Mean</td>
<td>3.8</td>
<td>5</td>
<td>0.0550</td>
<td>109</td>
<td>77.8</td>
<td>70.5</td>
<td>15.20</td>
<td>18.99</td>
</tr>
<tr>
<td>7 years</td>
<td>PPD</td>
<td>1</td>
<td>2</td>
<td>0.0209</td>
<td>1048.33</td>
<td>~</td>
<td>499.6</td>
<td>1118.5 fruits</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>13</td>
<td>0.0379</td>
<td>342.67</td>
<td>695.5</td>
<td>432.6</td>
<td>18.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>14</td>
<td>0.0195</td>
<td>715.33</td>
<td>702.11</td>
<td>551</td>
<td>21.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>22</td>
<td>0.0311</td>
<td>706.67</td>
<td>703.25</td>
<td>492.8</td>
<td>16.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>19</td>
<td>0.0657</td>
<td>289</td>
<td>620.4</td>
<td>400.6</td>
<td>24.90</td>
</tr>
<tr>
<td>Mean</td>
<td>18</td>
<td>5</td>
<td>0.0359</td>
<td>666.67</td>
<td>~</td>
<td>475.3</td>
<td>20.48</td>
<td>70.28</td>
</tr>
<tr>
<td>UP</td>
<td>2</td>
<td>26</td>
<td>0.0639</td>
<td>406.67</td>
<td>536.67</td>
<td>58.33</td>
<td>46.45</td>
<td>417.27</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>14</td>
<td>0.0312</td>
<td>447.33</td>
<td>506.89</td>
<td>76</td>
<td>36.78</td>
<td>65.16</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>18</td>
<td>0.0355</td>
<td>506.33</td>
<td>506.75</td>
<td>62.67</td>
<td>32.07</td>
<td>25.85</td>
</tr>
<tr>
<td>Mean</td>
<td>19.2</td>
<td>5</td>
<td>0.0445</td>
<td>314</td>
<td>468.2</td>
<td>69.67</td>
<td>8.54</td>
<td>9.59</td>
</tr>
<tr>
<td>10 years</td>
<td>PPD</td>
<td>1</td>
<td>15</td>
<td>0.0172</td>
<td>868.67</td>
<td>~</td>
<td>285.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>23</td>
<td>0.0185</td>
<td>1237.67</td>
<td>1053.17</td>
<td>318</td>
<td>45.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>20</td>
<td>0.0294</td>
<td>679.33</td>
<td>928.55</td>
<td>427</td>
<td>37.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>23</td>
<td>0.0380</td>
<td>605</td>
<td>847.66</td>
<td>235</td>
<td>30.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>18</td>
<td>0.0241</td>
<td>745.67</td>
<td>827.26</td>
<td>361.2</td>
<td>26.93</td>
</tr>
<tr>
<td>Mean</td>
<td>19.8</td>
<td>5</td>
<td>0.0322</td>
<td>589</td>
<td>~</td>
<td>325.36</td>
<td>35.03</td>
<td>139.47</td>
</tr>
<tr>
<td>UP</td>
<td>1</td>
<td>19</td>
<td>0.0322</td>
<td>589</td>
<td>~</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>22</td>
<td>0.0233</td>
<td>941</td>
<td>765</td>
<td>69</td>
<td>18.35</td>
<td>164.84</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>14</td>
<td>0.0353</td>
<td>396.33</td>
<td>642.1</td>
<td>63</td>
<td>21.30</td>
<td>58.61</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>18</td>
<td>0.0341</td>
<td>526.33</td>
<td>613.15</td>
<td>69.3</td>
<td>53.39</td>
<td>80.60</td>
</tr>
<tr>
<td>Mean</td>
<td>18.6</td>
<td>5</td>
<td>0.0308</td>
<td>648.33</td>
<td>620.18</td>
<td>56.3</td>
<td>4.89</td>
<td>6.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>20</td>
<td>0.0308</td>
<td>648.33</td>
<td>620.18</td>
<td>56.3</td>
<td>4.89</td>
</tr>
</tbody>
</table>

UP = uniform probability; PPD = probability proportional to the diameter; m = number of paths taken or sampling intensity; Fruits/branch = Number of fruits in the terminal branch; cvy = estimate of the coefficient of variation; țiQ = average number of fruits; Er = Relative sample error; țiQm = estimate of the number of fruits of the tree given by the path m; Qmi = unconditional probability of the selection of the branch i in the path m.


Mokochinski et al.
CANCINO & SABOROWSKI (2007) stated that the RBS provided estimates with minor errors when used in trees with regular crowns, assuming that the possible paths are equal in size. According to these authors, trees with irregular crowns and different sizes may present larger errors as a result of the probability of selecting branches between the paths.

LAMIEN et al. (2007), analyzing the production of Shea (Vitellaria paradoxa) in Africa, reported that fruit production occurs unevenly on trees with cross-pollination and that such differences were the result of the wind. The pecan tree also showed differences in the number of fruits in the terminal branches, contributing to the increase of errors in sampling.

WILLIAM (1989), estimating biomass of Pinus sp. with the branch sampling technique, described that the error tends to decrease with increasing sample size. BORGES (2009), estimating the production of pequi fruit with RBS using 10 paths, reported that the two sampling techniques (PPD and UP) underestimated the real fruit production, which occurred due to the great variability of fruit distribution inside the crown.

The relative efficiency resulted in mean values above 100% (588.5% - 5 years; 641.7% - 7 years; 1,120.7% - 10 years), indicating that the UP was more efficient than the PPD. Range inoefficients of variation (cv) and the time differences became key factors in analyzing the efficiency of the techniques. For the PPD, the difficulty of measuring the diameter near the node, makes the technique less efficient.

The decision to determine which sampling technique is more efficient for quantifying an attribute of interest must take into account the time spent on sampling and the coefficients of variation. The relative efficiency is not necessarily an indication that the estimate is close to the real value obtained from the census. The value of the relative efficiency is the basis for choosing which technique is more efficient.

For trees at 5 and 7 years, the relative efficiency values were similar. The highest value was reported for trees at 10 years. Due to the increased number of nodes in larger trees, there is an increase in the sampling time, resulting in an increase in the relative efficiency.

Since time is a factor that affects the relative efficiency, when repeated paths are sampled again, the time is reduced. The sampling time for pecan trees was important in the efficiency of the UP technique because the measurement of diameters

was not required, making this technique faster and affecting its relative efficiency.

CONCLUSION

Random sampling of branches did not prove efficient enough to estimate the production of pecan fruit in the evaluated ages (5, 7 and 10 years). The two evaluated techniques (uniform probability and probability proportional to the diameter) showed similar estimates of production and similar experimental errors.

ACKNOWLEDGEMENTS

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support.

REFERENCES

BARBEIRO, L. S. Sistema de amostragem para quantificar a produção de sementes de Bertholletia excelsa h. b. k (Castanha do Brasil) na região de Orixinimá - PA. 2012. 130f. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal do Paraná, Curitiba, PR.

BORGES, L. M. Amostragem aleatória de ramos como técnica para quantificar a produção de frutos de Caryocar brasiliense camb. (caryocaraceae). 2009. 168f. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal do Brasília, Brasília, DF.


LÔBO, R. N. Método de Amostragem Aleatória de Ramos (AAR) como técnica para estimar volume, biomassa e estoque em cerrado sensu-stricto. 2009. 80f. Monografia (Graduação em Engenharia Florestal) - Universidade de Brasília, Brasília, DF.

LÔBO, R. N. Amostragem aleatória de ramos como método para estimar os estoques em volume, biomassa e carbono em Floresta Estacional no estado de Goiás. 2012. 127f. Dissertação (Mestrado em Ciências Florestais) - Universidade de Brasília, Brasília, DF.
