Estimating Apple Fruit Size Distribution from Early-season Fruit Diameter Measurements

Richard P. Marini
Department of Plant Science, The Pennsylvania State University, 203 Tyson Building, University Park, PA 16802

James R. Schupp
Fruit Research and Extension Center, The Pennsylvania State University, 290 University Drive, Biglerville, PA 17307

Tara Auxt Baugher
The Pennsylvania State University, Cooperative Extension in Adams County, 670 Old Harrisburg Road, Gettysburg, PA 17325

Robert Crassweller
Department of Plant Science, The Pennsylvania State University, 7 Tyson Building, University Park, PA 16802

Abstract. In three experiments, diameters of apples representing 7% to 30% of the fruit on a tree were measured at ~60 days after full bloom. Using previously published regression equations, the early-season fruit diameter values were used to estimate apple fruit weight at harvest (FWH). At harvest, all fruit on sample trees were weighed and the distributions of estimated FWH for fruit measured early in the season were compared with distributions of the actual FWH for whole trees. Actual FWH was normally distributed for only one of the three experiments. Although the estimated mean FWH averaged for the 10 trees was within 9% of the actual mean FWH for all three experiments, the distribution of estimated FWH differed significantly from the actual distribution for all three experiments. All fruit were then assigned to appropriate commercial fruit sizes or box counts (number of fruit/19.05 kg). Fruit size tended to peak on the same four box counts for the estimated and actual populations, but the estimated populations had too few fruits in the small- and large-size box counts. Using early-season estimates of FWH, commercial apple growers and packers can predict fairly accurately the percentage of the crop that will fall into the peak box counts, but a more accurate early-season estimate of the fruit size distribution will likely require measuring 50% of the fruit on a tree.

Accurate early-season estimates of yield and fruit size distribution would aid apple growers and packers to estimate harvest labor, the number of bins, and the amount of storage space needed for the upcoming harvest. Accurate estimates of fruit size would also allow packers to better negotiate with potential buyers and retailers to develop marketing plans for the upcoming crop. There is currently no method to estimate packout at harvest from early-season fruit size estimates. The first step in estimating fruit size distribution requires models to estimate fruit size at harvest from early-season measurements. Apple fruit diameter at harvest was reasonably well predicted from early-season fruit diameter measurements for orchards within fairly close proximity (Batjer et al., 1957; Forshey, 1971). Marini et al. (2019a) recently reported that fruit diameter and fruit weight at harvest could be estimated from fruit diameter measurements at 60 d after full bloom (DAFB) for three cultivars growing in three fruit-growing regions of Pennsylvania, and coefficients of determination associated with the linear models ranged from 0.54 to 0.96. Adding the ratio of (fruit weight/fruit diameter) at 60 DAFB to the model with fruit diameter at 60 DAFB, produced less biased models with improved coefficients of determination, and predicted values were more similar to the observed values. The fruit weight/diameter ratio was positively related to cumulative growing degree days for the 60 d before the fruit were measured and tended to be lower in years when fruits were exposed to frosts.

The second step in estimating fruit size distribution for an individual tree involves sampling a tree adequately to obtain an unbiased estimate of early-season fruit size that represents the fruit size distribution for the whole tree. To assess the effect of various treatments on average fruit size on trees, researchers typically measured 5 to 35 fruit per tree (Byers and Carbaugh, 1981; Dozier et al., 1980; Greene, 1986; Miller, 1982; Schupp et al., 2002). Marini (2001) reported that estimates of mean fruit weight from a 20-fruit sample per tree differed from the true mean value by ≈13%. Estimates obtained by weighing all fruit on three limbs per tree differed from the true value by 11% to 19%. He suggested that sampling ≈25% of the fruit on a tree may be necessary to estimate mean fruit weight accurately. Over the past 15 years, many semidwarf orchards with relatively large trees have been replaced with more intensive orchards with smaller tree canopies. Smaller canopies have less variable light distribution, and fruit size may also be less variable. To determine the effect of canopy size on fruit size distribution, Marini et al. (2019b) divided vertical axis trees into eight vertical wedges or slices, and at harvest all fruit within each section were weighed. Fruit from most canopy sections were normally distributed, and distributions were similar for most sections. Therefore, we should be able to estimate fruit size distribution for a tree at harvest by measuring all fruit from a vertical section of tree representing at least 12.5% of the canopy.

Although Batjer et al. (1957) used early-season fruit diameter measurements to develop tables of estimated fruit diameter at harvest, they did not verify the models by comparing predicted fruit diameter to actual fruit diameter. They also did not consider how a tree should be sampled to obtain estimates that are representative of the whole tree. To our knowledge, no one has attempted to estimate the size distribution of a small population of trees at harvest from a sample of fruit measured early in the season. The final step in developing a protocol for estimating fruit size distribution for a group of trees at harvest from early-season fruit diameter measurements is to compare predicted fruit size distributions obtained from a sample of fruit with the actual distribution obtained from the whole tree at harvest. Therefore, the objective of this study was to determine whether previously published sampling protocols and regression equations developed with data from several orchards and several years (Marini et al., 2019a, 2019b) can be used to predict distributions of fruit weight and box sizes for a small population of trees. To test this hypothesis, distributions of fruit weight estimated by measuring diameters of all the fruit developing in about 12.5% or 25% of the volume of
the canopy at 60 DAFB were compared with the actual distributions obtained by weighing all fruit on the tree at harvest.

Materials and Methods

General. Trees from three apple plantings in two orchards were used for this study. In 2016, 10 ‘Buckeye Gala’ trees on M.9 NAKBT337 rootstock, planted in 2008, were selected at Penn State’s Fruit Research and Extension Center (FREC) in Biglerville, PA (lat. 39°56’22”N, long. 77°15’25”W). Trees were spaced 0.91 × 3.7 m, were ≥3 m tall and 1.8 m in diameter at the base of the canopy, trunk cross-sectional area (TCSA) averaged 3.8 cm², and rows were oriented northeast – southwest. In 2017, 10 ‘Buckeye Gala’ and 10 ‘Aztec Fuji’ trees on M.9 NAKBT337 rootstock, planted in 2009 and trained to a central axis, were selected in a commercial orchard in Fishertown, PA (lat. 40°7’40”N, long. 78°31’37”W). Trees were planted 0.91 × 3.7 m, were about 2.4 m tall and 1.8 m in diameter at the base of the canopy, average TCSA was 2.4 cm² for ‘Gala’ and 2.6 cm² for ‘Fuji’ trees, and rows were oriented in a northwest–southeast direction. For all three experiments, trees with typical crop loads were selected randomly from two adjacent rows. Trees for all three experiments were thinned with 600 mg L⁻¹ carbaryl (Carbaryl 4L; Loveland Products, Inc., Greeley, CO) plus 2.5 mL L⁻¹ SunSpray Ultra-Fine spray oil (HollyFrontier Refining and Marketing LLC, Plymouth, PA) delivered with an air-blast sprayer calibrated to apply 935 L ha⁻¹, and follow-up hand thinning was required to retain one fruit per spur every ≈15 cm along the limb. Trees at the FREC had drip irrigation, and trees at Fishertown were not irrigated.

Fruit sampling differed slightly for the two orchard locations. For trees at the FREC, each tree was randomly assigned to a section of the canopy representing approximately one-eighth of the canopy (north, northeast, east, southeast, south, southwest, west, or northwest). At 60 DAFB, two strips of flagging tape were placed vertically on the periphery of each canopy to isolate that section visually from the remainder of the canopy. The strips of flagging tape delineated a three-dimensional vertical wedge of canopy from the top to the bottom of the canopy and from the canopy periphery to the trunk (Fig. 1). Trees at Fishertown were randomly assigned to a section of the tree representing ≈25% of the tree (north, south, east, or west) that was delineated by placing two strips of flagging tape vertically in the tree. At ≈60 DAFB and after hand-thinning, the diameter of every fruit within the vertical wedge of the canopy was measured with an electronic fruit size measurer (EFM; QA Supplies LLC, Norfolk, VA 23502). Fruits outside of the marked wedge were not measured. At commercial harvest, all fruits on each tree were harvested and weighed. Fruit harvested from the vertical wedge of canopy were not separated from fruit from the rest of the tree, and the weight of each individual fruit on each tree was recorded. Hereafter, fruit within the vertical wedge measured at 60 DAFB and used to predict fruit weight at harvest are referred to as the “estimating sample.” The “actual population” consisted of all fruit on a tree (fruit from the estimating sample plus those not in the estimating sample). Fruit at the FREC were weighed on an electronic single-lane fruit sizer equipped with a digital load-cell (Durand-Wayland, Inc., LaGange, GA). Each fruit harvested at Fishertown was weighed on an Ohaus Scout Pro portable balance (Mettler Toledo, Columbus, OH) to the nearest gram.

Predicting fruit weight. Methods for predicting fruit weight at harvest differed for the two cultivars. Multiple regression models for predicting FW from early-season fruit diameter measurements were previously reported (Marini et al., 2019a). For ‘Gala’, predicting FW at harvest involved two steps. First, fruit diameter at harvest (FDH) was estimated with a multiple linear regression model, where fruit diameter at 60 DAFB (FD60) and the ratio of (FW60/FD60) were included as predictor variables (Fig. 4 in Marini et al., 2019a). Then the predicted fruit weight at harvest (EFWH) was obtained with a previously published quadratic model in which predicted FDEH (EFDH) was the predictor variable (Fig. 1 in Marini et al., 2019a). For ‘Fuji’, EFWH was obtained directly from FD60 and the ratio of (FW60/FD60) (Fig. 7 in Marini et al., 2019a):

\[
\text{'Gala': } EFDH (g) = 85.39 + (0.241' FD60) - [52.955' (FW60/FD60)] + [0.586' (FW60/FD60)]
\]

\[ R^2 = 0.627, \text{where (FW60/FD60) = 0.992} \]

\[
\text{'Fuji': } EFWH (g) = 17.03 - (1.48'EFDH) + [0.046'EFDH] \]; \quad R^2 = 0.972
\]

\[
\text{where (FW60/FD60) = 0.78.}
\]

Statistical analyses. Descriptive statistics for each data set were obtained with SAS’s PROC MEANS (Table 1). Analysis of variance (ANOVA) was used to determine whether the mean FWH for the estimating sample was equal to the actual mean FWH for whole trees. Before performing the ANOVA for each data set, homogeneity of variances was evaluated with Levene’s test by performing an ANOVA on the absolute values of the residuals with PROC GLM (Littell et al., 2002). The assumption of homogenous variances was rejected for ‘Gala’ and ‘Fuji’ at Fishertown, so heterogeneous models were fit with PROC MIXED to compare the estimated mean FWH with the actual mean FWH. Proc UNIVARIATE was used to test the hypothesis that populations of estimated and actual FWH were normally distributed; test statistics and P values for the Kolmogorov-Smirnov (KS), Shapiro-Wilk (W), Cramer-von Mises (CVM), and Anderson-Darling (AD) tests are presented in Table 2. The empirical distribution function (EDF) is a nonparametric estimate of the cumulative distribution function and was used to describe the populations of estimated and actual FWH. The value of the EDF at a given point is equal to the proportion of observations from the sample that are less than or equal to that point. The nonparametric KS two-sample test, obtained with SAS’s PROC NPAR1WAY, was used to test equality of EDFs for estimated and actual FWH.

Results

The number of fruits harvested per tree, along with the minimum, mean, and maximum FWH plus standard deviations for each tree are presented in Table 1. Previously, we found that sampling all the fruit from a vertical wedge of canopy representing ≈12.5% of the canopy volume provided a distribution of fruit sizes similar to the distribution for all fruit on a tree (Marini et al., 2019b). Following this sampling scheme at 60 DAFB for ‘Gala’ trees at the FREC, we measured fruit diameters of all fruits in a vertical wedge containing 12.5% of the canopy. Because fruit are not uniformly distributed throughout the canopy, these wedges contained 7% to 14% of the fruit on a tree. In 2017, the percentage of canopy sampled at Fishertown was increased to 25%, and the estimating sample contained ≈25% to 30% of the fruit on the tree. The mean FWH for the estimating sample differed significantly from the actual mean FWH for two of the three experiments. When averaged over the 10 trees, FWH for the estimating sample was within 10% of the actual mean. For ‘Gala’ in Fishertown, the mean FWH for the estimating sample was equal to the actual mean FWH (P = 0.94).

‘Gala’ at FREC in 2016. On the basis of the four tests for normality, both the
estimating sample and the actual population for FWH deviated significantly from normality (P < 0.05; Table 2). Histograms provide a general idea of the shape of a distribution, but skewness and kurtosis provide more precise numerical measures of shape. Lack of symmetry (skewness) and pointiness (kurtosis) are two ways in which a distribution can deviate from normal (Ghasemi and Zahedias, 2012). Skewness indicates the amount and direction of skew (departure from horizontal symmetry), and kurtosis indicates how tall and sharp the central peak is, relative to a standard bell curve. Both distributions were about symmetrical because skewness values were less than |0.5|. The negative skewness statistic for the estimated population indicated that the distribution was slightly skewed to the left (small fruit), and the actual population was skewed to the right (large fruit); kurtosis was greatest for the whole-tree population. The estimating sample had fewer extreme values as indicated by the smaller range of FWH values (Table 1) and the narrower base on the histogram (Fig. 2). Plots of the EDFs for FWH indicate that FWH estimated from early-season fruit diameter tended to underestimate the proportion of fruit smaller than \( \approx 140 \) g and overestimated the proportion of fruit greater than 150 g.

Table 1. Number of fruit and minimum, mean, and maximum fruit weights (FW), plus the standard deviation (SD) for fruit harvested from the whole tree and for fruit from the estimating sample for 10 trees in three orchards.

| Tree | Estimating sample FW at harvest (g) | Actual FW for whole-tree at harvest (g) |
|------|-----------------------------------|----------------------------------------|
|      | N  Min. Mean Max. so              | N  Min. Mean Max. so                    |
| 'Gala' at FREC in 2016 | 1 18 134 166 187 14.5 | 129 82 196 248.8 |
|      | 2 15 142 153 177 10.9           | 130 64 170 247 34.6  |
|      | 3 16 118 159 182 18.4            | 125 75 169 278 31.4  |
|      | 4 16 111 151 173 15.3            | 130 76 166 260 34.6  |
|      | 5 12 93 140 168 24.2            | 160 68 161 290 35.1  |
|      | 6 20 134 149 182 11.9           | 184 57 158 243 31.9  |
|      | 7 24 107 142 177 16.9           | 183 64 155 239 33.7  |
|      | 8 25 118 147 168 3.6            | 283 49 146 255 34.6  |
|      | 9 23 126 150 177 12.5           | 348 37 143 286 38.4  |
|      | 10 25 104 136 159 15.6          | 172 37 131 286 36.7  |

Mean ± 159.3 ± 5.3

'Gala' at Fishertown in 2017

Mean ± 132 ± 4.56

Mean ± 245 ± 3.73

P = 0.0011

Table 2. Values for four statistics used to test the normality of populations of fruit weight at harvest for the estimating sample and for the actual population for three experiments with 10 trees in each of three orchards.

| Normality test | Estimated from sample population | Actual whole-tree population |
|----------------|----------------------------------|-----------------------------|
| Test stat.     | P value                          | Skewness | Kurtosis | Test stat. | P value | Skewness | Kurtosis |
| SK             | 0.085 ≤0.010                    | -0.291  | 0.149    | 0.062 <0.010 | 0.107    | 0.357    |
| W              | 0.986 ≤0.012                    | —       | —        | 0.994 <0.001 | —        | —        |
| KVM            | 0.148 0.025                     | —       | —        | 0.758 <0.005 | —        | —        |
| AD             | 0.847 0.030                     | —       | —        | 4.054 <0.005 | —        | —        |

'Gala' at Fishertown in 2017

| SK             | 0.094 ≤0.010                    | 0.171  | 0.021    | 0.033 0.038 | 0.220 | -0.190 |
| W              | 0.988 0.034                     | —       | —        | 0.994 >0.006 | —        | —        |
| KVM            | 0.260 ≤0.005                    | —       | —        | 0.124 >0.055 | —        | —        |
| AD             | 1.358 ≤0.005                    | —       | —        | 0.851 0.030 | —        | —        |

'Fuji' at Fishertown in 2017

| SK             | 0.097 ≤0.010                    | -0.269 | -0.394 | 0.029 >0.150 | -0.020 | -0.302 |
| W              | 0.980 0.005                     | 0.997  | 0.209 | 0.070 >0.250 | —        | —        |
| KVM            | 0.249 ≤0.005                    | —       | —        | 0.478 0.240 | —        | —        |
| AD             | 1.471 ≤0.005                    | —       | —        | —        | —        | —        |

### Tests for normality: AD = Anderson-Darling test; KVM = Cramér-von Mises criterion; SK = Kolmogorov-Smirnov test; SW = Shapiro-Wilk test.
According to the KS two-sample test, the two EDFs were significantly different \( (P < 0.0001) \). The two EDFs differed due to the smaller range of FWH for the estimated population. The smallest fruit from the estimating sample was 93 g, whereas 5% of the actual population was less than 93 g. The largest single fruit from the estimating sample was 187 g, whereas 19% of the fruit in the actual population was greater than 187 g. The box count (number of fruit/19.05 kg) peaked on 125 and 113 for the estimated and actual populations, respectively (Fig. 2). Similar to FWH, the actual population had more fruit in the smallest and largest box sizes (Table 3). The EDFs were also significantly different (Fig. 2). The difference in EDFs was primarily due to the lack of small fruit in the estimated population because 17.7% and 46.7% of the estimated and actual population had large fruit with box counts less than 125, respectively (Table 3).

‘Gala’ at Fishertown were smaller than at Fishertown in 2017. ‘Gala’ fruit harvested at Fishertown were smaller than at
Table 3. Percentage of apple fruit in 12 commercial box sizes for the estimated values, obtained by predicting fruit weight at harvest from early-season fruit diameter measurements, and for the actual population obtained by weighing all fruit on the tree.

| Box count | ‘Gala’ at FREC 2016 | ‘Gala’ at Fishertown 2017 | ‘Fuji’ at Fishertown 2017 |
|-----------|---------------------|--------------------------|--------------------------|
|           | Estimated | Actual | Estimated | Actual | Estimated | Actual |
| 56        | 0.0       | 0.0    | 0.0       | 0.0    | 0.0       | 2.0    |
| 64        | 0.0       | 0.3    | 0.0       | 0.0    | 16.8      | 7.3    |
| 72        | 0.0       | 0.9    | 0.0       | 0.3    | 32.0      | 18.5   |
| 80        | 0.0       | 2.6    | 0.0       | 3.2    | 22.2      | 15.9   |
| 88        | 0.0       | 8.3    | 0.0       | 6.9    | 17.7      | 17.9   |
| 100       | 3.1       | 14.3   | 0.0       | 11.8   | 7.9       | 15.4   |
| 113       | 15.6      | 19.8   | 1.6       | 11.8   | 2.5       | 9.1    |
| 125       | 39.7      | 16.5   | 14.4      | 11.8   | 1.0       | 0.0    |
| 138       | 23.7      | 11.6   | 38.4      | 14.8   | 0.0       | 0.0    |
| 150       | 6.8       | 7.9    | 24.4      | 11.3   | 0.0       | 0.0    |
| 163       | 2.6       | 0.1    | 12.8      | 4.0    | 0.0       | 0.0    |
| 175+      | 4.7       | 17.2   | 8.4       | 35.9   | 0.0       | 14.0   |

*Number of fruits in a 19.05-kg carton.

FREC, probably because trees at FREC were trickle irrigated, but trees at Fishertown were not irrigated. Distributions for the estimating sample and the actual population significantly differed from normal. Both distributions were symmetrical with small values for skewness and kurtosis (Table 1), but the actual population had greater kurtosis with a higher peak near the midpoint (Fig. 3). Similar to the ‘Gala’ at FREC, the range of values was greatest for the actual population resulting in a wider base for the histogram. The plot of the EDFs showed that predicting FWH from early-season fruit diameter measurements tended to underestimate the proportion of fruit smaller than 130 g and overestimate the proportion of fruit greater than 130 g. The histograms for estimated and actual box counts were different, and both differed significantly from normal (Fig. 3). The estimating population had no fruit larger than 100 count, as opposed to 10.3% for the actual population (Table 3). The EDFs for the two populations also differed significantly (Fig. 3), and box count predicted from early-season measurements underestimated the percentage of large fruit in the 88- to 138-count categories. Approximately 50% of the fruit from both populations were small and fell into the 150+ count, but only 8% of the estimated fruit and 36% of the actual fruit fell into the 175+ box count and would likely be sold for juice.

‘Fuji’ at Fishertown in 2017. As in the other two experiments, the range of FWH was greater for the actual population than for the estimating sample (Table 1). Tests for normality indicated that the actual population, but not the estimated population approximated normality (Table 2). Both populations were symmetrical. The estimated population was slightly skewed to the left, but the actual population was not skewed. The estimated population also had higher kurtosis. FWH for the estimated population peaked between 240 g and 270 g and the actual population peaked between 210 g and 250 g (Fig. 4). The plot of EDFs showed that the estimated population underestimated the actual FWH over the entire range of FWHs except for large fruit greater than 280 g (Fig. 4). The histograms for box count were highly skewed to the left (large fruit) with skewness values of 1.13 and 0.948 for the estimated and actual populations, respectively. The histogram was more peaked for the estimated population (kurtosis = 1.51) than for the actual population (kurtosis = 0.517). The plot of the EDFs showed that the estimated box counts of 72 and 80 were overestimated for the estimating population.

Discussion

The sampling protocol used in this study produced accurate estimates of mean FWH, when averaged over the 10 trees, but the range of actual FWH for all 30 trees was underestimated in the estimating sample. When averaged over the 10 trees, for two of the three experiments, the mean FWH for the estimating sample was within 10% of the mean for the actual population, with similar standard errors. For ‘Gala’ at Fishertown in 2017, the mean FWH for the estimating sample was equal to the actual mean, and standard errors were similar. These results support our previous report (Marini et al., 2019b), in which fruit size measurements of 25% to 30% of the fruit on 10 or more trees in intensive orchards provided estimates of mean FWH equal to or better than previous sampling schemes at harvest for semidwarf trees (Marini, 2001). However, the estimating sample underestimated the number of small and large fruit on every tree, and the mean FWH for the estimating sample was within 15% of the actual mean for only 18 of the 30 trees in this study. For unknown reasons, the estimating samples of fruit in this study were less similar to the actual whole-tree populations than in our previous study (Marini et al., 2019b). Standard deviations for whole ‘Gala’ trees were approximately three times higher than for the estimating sample, and for ‘Fuji’, the standard deviations for the whole trees were at least 25% higher than for the estimating samples. The estimating sample did not reflect the fruit size distribution as accurately as we expected, so a greater proportion of the fruit on a tree is likely needed as an estimating sample. Because the estimating sample was composed of fruit close to the average size, mean FWH was estimated more accurately than the distribution of FWH.

Distributions for actual whole-tree FWH were expected to be normal due to the large sample sizes, but the actual populations of fruit were significantly different from normal in two of the three experiments. According to the central limit theorem, if samples are taken from a population with a normal distribution, as the sample size increases, the parameters of the distribution will approach normal. Because the estimating sample for the 10 trees had a total of only 194 to 250 fruit, we expected that it might not be reasonable to assume it to be normally distributed. However, we expected that the total of 728 to 1844 fruits harvested from the 10 trees for each experiment would be normally distributed. The departure from normality may be due to crop load management strategies of chemical thinning plus follow-up hand thinning. Altman and Bland (1995) showed that as the number of random samples increased from 100 to 500 the distribution parameters approached normal. They also showed that the range of values increased with increasing sample size, and this likely explains the absence of large and small FWH values in our estimated populations. We presented four tests for normality because statisticians do not agree on which is most appropriate and are influenced differently by sample size and by different characteristics of the population (Ghasemi and Zahedias, 2012; Le Boedec, 2016). The KS test has fallen out of favor for testing normality because it is sensitive to extreme values and has low power. The W test is based on the correlation between data and the corresponding normal scores and is more powerful than KS. The CVM criterion is used to judge the goodness-of-fit of a cumulative distribution function compared with a given EDF (one-sample test) and can be extended to compare two EDFs (two-sample test). The AD test is a test of whether a given sample of data is drawn from a given probability distribution. It can be used to test whether a normal distribution adequately describes a set of data and is one of the most powerful tests for detecting most departures from normality. Ahad et al. (2011) prefer the W test because it rejects the null hypothesis.
of normality test at the smallest sample size compared with other tests, for all levels of skewness and kurtosis of these distributions. Although the $P$ values for the four tests did not always agree, in general, $P$ values for both estimated and actual FWH populations were less than 0.05, indicating that the data were not from a normal distribution. Therefore, it cannot be assumed that the actual population of fruit weights in an orchard is normally distributed.

To our knowledge, this is the first report in which distributions for fruit size were formally tested for normality. Histograms of nectarine fruit size were reported by Naor et al. (1999), but normality and equality of distributions were not tested. Lötzé and Bergh (2004) developed multiple regression models to estimate FWH for ‘Golden Delicious’ apple. Variables in the model included average FW at 42 DAFB for all fruit.

Fig. 3. Histograms of fruit weight (top) and box count (middle) with normal curves for estimated and actual values for ‘Gala’ trees at Fishertown in 2017. Empirical cumulative distribution functions for fruit weight (bottom left) and box counts (bottom right) for estimated and actual values, where the solid line represents the actual values. The Kolmogorov-Smirnov test statistics ($K_{sa}$) are significant, indicating that the two distribution functions are different.
on sample trees, the percentage of fruit in different size groups at 42 DAFB, crop density, and time of fruit thinning, and the coefficient of determination was 0.8. However, model selection criteria were not reported, and some of the variables in the model were likely collinear. Collinearity may cause imprecise estimates of coefficient values, the standard errors may be large, and the coefficient of determination may be inflated (Myers, 1990). In addition, equality of their estimated and actual FWH distributions was not formally tested (Lütze and Bergh, 2004). Webb et al. (1980) weighed all fruit on ‘Golden Delicious’ trees and stated that “the distribution of fruit weights on the 8 trees has been shown to fit the normal curve with common variance,” but histograms of FWH were not reported and normality was not formally tested. Although histograms of fruit size often appear to represent a normal distribution and histograms of FWH estimated by various methods may appear to be similar, our analyses indicate that apple FW may not be normally distributed, and histograms of distributions that appear similar may be significantly different.

Fig. 4. Histograms of fruit weight (top) and box count (middle) with normal curves for estimated and actual values for ‘Fuji’ trees at Fishertown in 2017. Empirical cumulative distribution functions for fruit weight (bottom left) and box counts (bottom right) for estimated and actual values, where the solid line represents the actual values. The Kolmogorov-Smirnov test statistics (Ksa) are significant, indicating that the two distribution functions are different.
The distributions for actual FW are often of interest to researchers, but growers and packers are more interested in box count because apples are sold based on the number of fruits in a 19.05 kg carton. For all three experiments, the distributions of estimated and actual box counts were significantly different (Fig. 2–4). However, fruits smaller than 165-count will typically be sold for juice and for two of the three experiments, the percentage of fruit smaller than 165-count was similar (Table 3). For all three experiments, the estimated populations underpredicted the larger fruit sizes, but less than 10% of the crop usually fell into these large sizes. It is unreasonable to expect perfect estimates of fruit size, but packers and growers are probably interested to know into which of the three or four box sizes most fruit will fall. The four most common size categories for ‘Gala’ at the FREC in 2016 included 100-, 113-, 125-, and 138-counts, and the combined percentage of fruit in those four sizes were 86% and 63% for estimated and actual populations, respectively. For ‘Gala’ at Fisherton in 2017, the most common fruit sizes were 113-, 125-, 138, and 150-count and included 79% and 50% of the fruit for the estimated and actual populations, respectively. For ‘Fuji’ at Fisherton in 2017, the most common fruit sizes included 72-, 80-, 88-, and 100-count, and 80% and 68% of the estimated and actual populations fell into those four categories. If more accurate estimates are desired, then a larger estimating sample may be necessary, possibly by obtaining early-season FW estimates from 50% of the fruit on the sample trees.

Accurate predictions for packout depend on early-season estimates that accurately reflect the actual whole-tree population, as well as reliable models for estimating FW. Although our previous work indicated that measuring 25% of the fruit on a tree provided good estimates of fruit size distributions, our current study suggests that a larger estimating sample is preferable. The relationship between early-season fruit size and fruit size at harvest varies with season and orchards within season. The models used in this study were expected to be fairly robust because they were developed by combining data from more than a single combination of year and orchard where the regression models were similar, and coefficients of determination ranged from 0.54 to 0.63 (Marini et al., 2019a). Previously published coefficients of determination from similar studies with other cultivars ranged from 0.21 to 0.67 (Batjer et al., 1957; Forshey, 1971; Williams et al., 1969). Because all fruits on a tree do not grow at the same rate from 60 DAFB to harvest, it is unlikely that early-season packout estimates can be much more accurate than those presented for our study unless the number of fruit for the estimating sample is increased. Another way to improve the accuracy of fruit size estimates is to measure fruit later in the season because Forshey (1971) reported that the correlation between fruit diameter before harvest and at harvest improved as measurements were made closer to harvest.

All trees in this study were trained to the vertical axis system and had a conical shape, but trees at the FREC were larger than those at Fisherton. Fruit distribution and size may be different for larger trees or trees trained to a different system where there is more shade in the lower canopy. Therefore, results from this study may not be extrapolated to trees growing in different orchard systems.

Previous researchers simply correlated fruit diameter at harvest with early-season fruit diameter measurements. This study is the first attempt to predict fruit weight and packout distributions for a small population of trees. Although the percentage of fruit in the smaller and larger box sizes were underestimated, we were able to identify the predominant box sizes. Information from this study can be used by growers and packers to obtain reasonable estimates of fruit size distributions for a small population of trees. Additional research is needed to determine how to sample a block of trees to obtain estimates for the block that accounts for tree-to-tree variation. Agricultural engineers are working with pomologists to develop machine vision applications to estimate the number and size of fruit in intensive orchards. These automated systems may be able to measure fruit size early in the season, and using predictive equations such as ours, they may be able to predict yield and packout for an orchard.

**Literature Cited**

Ahad, N.A., T.S. Yin, A.R. Othman, and C.R. Yaacob. 2011. Sensitivity of normality tests to non-normal data. Sains Malays. 40:637–641.

Altman, D.G. and J.M. Bland. 1995. The normal distribution. BMJ 310:298.

Batjer, L.P., H.D. Billingsley, M.N. Westwood, and B.L. Rogers. 1957. Predicting harvest size of apples at different times during the growing season. Proc. Amer Soc. Hort. Sci. 70:46–57.

Byers, R.E. and D.H. Carbaugh. 1991. Effect of chemical thinning sprays on apple fruit set. HortTechnology 1:41–48.

Dozier, W.A., Jr., W.A. Griffy, and H.F. Burgess. 1980. Effect of growth regulators on the development of ‘Delicious’ apples. HortScience 15:743–744.

Forshey, C.G. 1971. Predicting harvest size of McIntosh apple. N.Y. Food and Life Sci. Bull. No. 9.

Ghasemi, A. and S. Zahedias. 2012. Normality tests for statistical analysis: A guide for non-statisticians. Intl. J. Endocrinol. Metab. 10:468–489.

Greene, D.W. 1986. Effect of paclobutrazol and analogs on growth, yield, fruit quality, and storage potential of ‘Delicious’ apples. J. Amer. Soc. Hort. Sci. 111:328–332.

Le Boedec, K. 2016. Sensitivity and specificity of normality tests and consequences on reference interval accuracy at small sample size: A computer-simulation study. Vet. Clin. Pathol. 45:648–656.

Littell, R.C., W.W. Stroup, and R.J. Freund. 2002. SAS for linear models. 4th ed. SAS Institute Inc., Cary, NC.

Lütze, E. and O. Bergh. 2004. Early prediction of harvest fruit size distribution of an apple and pear cultivar. Scientia Hort. 101:281–290.

Marini, R.P. 2001. Estimating mean fruit weight and mean fruit value for apple trees: Comparison of two sampling methods with the true mean. J. Amer. Soc. Hort. Sci. 126:503–510.

Marini, R.P., J.R. Schupp, T.A. Baugher, and R. Crassweller. 2019a. Relationships between fruit weight and diameter at 60 days after bloom and at harvest for three apple cultivars. HortScience 54:86–91.

Marini, R.P., J.R. Schupp, T.A. Baugher, and R. Crassweller. 2019b. Sampling apple trees to accurately estimate mean fruit weight and fruit size distribution. HortScience 54:1017–1022.

Miller, S.S. 1982. Regrowth, flowering, and fruit quality of ‘Delicious’ apple trees as influenced by summer pruning. J. Amer. Soc. Hort. Sci. 107:975–978.

Myers, R.H. 1990. Classical and modern regression with applications. 2nd ed. PWS-Kent, Boston, MA.

Naor, A., I. Klein, H. Hupert, Y. Grinblat, M. Pers, and A. Kaufman. 1999. Water stress and crop level interactions in relationship to nectarine yield, fruit size distribution, and water poten-

tials. J. Amer. Soc. Hort. Sci. 124:189–193.

Schupp, J.R., E. Fallahi, and I.J. Chun. 2002. Effect of chemical thinning sprays on apple fruit set. Scientia Hort. 94:596–598.

Williams, M.W., H.D. Billingsley, and L.P. Batjer. 1969. Early season harvest size prediction of ‘Bartlett’ pears. J. Amer. Soc. Hort. Sci. 94:596–598.