

A *Vitis vinifera* Cluster's Wing-related Structural Characteristics and Their Associations with Yield and Berry Composition

Michael E. Tarter

Division of Biostatistics, School of Public Health, University Hall, University of California, Berkeley, CA 94705

Stefano Poni¹

Istituto di Frutti-Viticultura, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29100 Piacenza, Italy

Additional index words. cordon, Fisher's *z* transformation, Gaussian quadrature, proximal branching, quartile, random sampling, vineyard row

Abstract. The hypotheses considered in this article concern the basic question, besides bearing a wing, in what ways do wing-bearing and non-wing-bearing clusters differ? Vines sampled at midseason were again selected at harvest. Each weight of a *Vitis vinifera* cluster sampled at midseason was multiplied by the number of clusters on the vine from which the cluster had been selected. Correlation coefficients between this quantity and the sampled vine's yield at harvest differed significantly in the sense that coefficients determined solely from the subset of sampled clusters on which a wing (a lateral arm originating from the peduncle and separate from the main body of the cluster) was present were found to be larger than coefficients determined from all sampled clusters. To shed light on distinguishing characteristics of clusters that bore wings, the weights of clusters that had been sampled at midseason were studied. Despite being weighed after the removal of their wings, clusters that had wings were found to be significantly heavier than clusters (sampled at the same midseason date) that had never had wings. Box and whisker plots were constructed to assess this finding as well to study the relationships between a *Vitis vinifera* rachis' (a cluster's principal axis) weight, length, and diameter and wing absence or presence. For each of the five vineyard blocks that we studied, the median rachis midseason diameters of wing-bearing clusters exceeded the median rachis diameters of non-wing-bearing clusters. Concerning 'Cabernet Sauvignon' *Vitis vinifera* clusters that had wings, it was also found that the late-season differences between the median soluble solids concentrations (°Brix) of wing-borne berries and the median °Brix of non-wing-borne berries were inappreciable.

The variate "cluster weight" is an important *Vitis vinifera* vine yield component and its main subcomponents are berry number and berry weight (Clingeleffer et al., 2000). This variate's values can be affected by factors that include scion/clonal variability, number of clusters grown on a shoot (Pratt, 1971; Tarter and Keuter, 2008), degree of branching on a cluster (Clingeleffer et al., 2000), and weather course at resumption of bud growth in spring (Carmo Vasconcelos et al., 2009). According to Dunn and Martin (2007), "A main axis, the rachis, carries primary side branches, which are also branched and look like small bunches. The nearest (proximal) branch to the base of the stalk is sometimes called an 'outer arm' or 'wing,' and can be highly variable in form, ranging from a bunch-like structure rivaling the main axis in size, all the way down through

smaller sizes to just a tendril." For the purposes of our studies, wings were defined as lateral arms originating from the peduncle (separate from the main body of the cluster).

Tarter and Keuter (2008) implemented a random-systematic-random approach, a composite procedure in which, as a first step, rows as well as vines within selected rows were selected randomly without replacement. This was followed by the systematic selection of shoots and, as a third step, clusters from each of these selected shoots were chosen randomly without replacement. Access to data obtained in this way facilitated new studies that concerned the independent variable, wing absence or presence. Accordingly, the hypotheses considered in this article concern the basic question, besides bearing a wing, in what ways do wing-bearing and non-wing-bearing clusters differ?

To shed light on the concomitants of wing presence, we sought to: 1) determine the degree to which wing presence or absence on a cluster is associated with a vine's yield at harvest; 2) establish if wing presence or absence correlates with other structural clus-

ter characteristics (i.e., rachis weight, length and diameter) and/or is influenced by a *Vitis vinifera* shoot's position along a cordon (cordon pruning leaves a permanent horizontal extension of the vine's trunk in place year after year), or a row's position within a vineyard; and 3) compare the final soluble solids concentrations (°Brix) of wings with clusters' °Brix (wing berries excluded). The latter hypotheses concern comparisons between wing and cluster °Brix similar to hypotheses concerning a berry's location along a cluster's rachis that were considered by Tarter and Keuter (2005).

Methods

Sample collection. Tarter and Keuter (2008) describe how 92 cordon pruned vines had been selected from each of five blocks of vines. These 460 sampled vines, all located in the Napa Valley, were grown on two *Vitis vinifera* 'Cabernet Sauvignon' blocks designated as CS1 and CS2, two 'Merlot' blocks, ME1 and ME2, and a 'Chardonnay' block, CH. All blocks contained spur-pruned cordon vines trained to a vertically shoot positioned system. Pre-harvest collection dates (Stage 35 according to Eichhorn and Lorentz classification as cited by Coombe, 1995) were: 21 and 22 July for CS1; 1, 7, and 8 Aug. for CS2; 3 and 5 Aug. for ME1; 24 and 26 Aug. for ME2; and 17 to 19 Aug. for CH. Dates chosen for midseason cluster removal were selected to satisfy two conditions. The first was to select clusters from CS1 and CS2 when, in the light of past experience, these clusters' weights could be expected to assume values approximately equal to half of what they would have been at harvest. The second condition was to take midseason measurements from all five blocks in as short a time interval as was practical.

The vines in CS1, ME1, and CS2 had bilateral cordons, and ME2 as well as CH had unilateral cordons. ME2 was planted in 1995 with 1.54 m by 2.13-m spacing with 121 rows oriented northeast/southwest, each with between 12 and 131 vines (9989 vines total). Vines were on 110R rootstock. CS2 was planted in 2001 with 1.54 m by 1.83-m spacing and 89 rows oriented northeast/southwest, each with between 87 and 94 vines (8830 vines total). CH was planted in 1991 with 1.54 m by 2.13-m spacing and 82 rows oriented east/west, each with between four and 162 vines (9596 vines total). The rootstock for both CS2 and CH was 101-14.

A random-systematic-random sampling method had helped to account for non-rectangular block geometry to decrease short-row bias when selecting rows within a block randomly and then choosing pairs of vines within the selected rows randomly without replacement (Tarter and Keuter, 2008). Weight and other measurements could only be made once clusters had been removed from the shoot to which they had been attached. This constraint motivated the sampling strategy, depicted by Figure 1, in which pairs of adjacent vines, not individual vines, were selected. Once the vine

Received for publication 26 Mar. 2010. Accepted for publication 14 June 2010.

¹To whom reprint requests should be addressed; e-mail stefano.poni@unicatt.it.

pair was selected, a marked elastic ribbon was used to systematically select four shoots, two on each vine pair member. Care was taken to guarantee that if, for example, the two shoots labeled A and C in Figure 1 were chosen on the leftmost of the two vines, then shoots labeled B and D were chosen from the rightmost of the two vines. This strategy tended to reduce the effect that midseason cluster removal had on a sampled vine's yield at harvest while simultaneously reducing the effort needed to gather information (because four, not a single, cluster could be collected from each specific sampled location within a block of vines). By choosing a permutation of the letters A, B, C, and D randomly (without replacement), the vine pair and marked ribbon-based method tended to reduce any effect that position along the cordon might have on variables such as cluster weight and °Brix.

To improve the sampling method, each elastic ribbon, data-gathering tool had been pre-marked on the basis of values often referred to as Gaussian nodes. Applications of this approach to sampling had been discussed by Tarter and Keuter (2005) in a vineyard context. To study wing-related structural characteristics, ribbons were marked at positions A, B, C, and D using tabled node values computed as described by Stroud and Secrest (1966). The optimal spacing of dependent variable values over which polynomial and other regression functions can be estimated has been extensively analyzed. Based on the findings of Hoel (1958) and Hoel and Levine (1964), these values have been found to be either the roots of the polynomials or derivatives of the same polynomials whose roots are numerically equal to the Gaussian nodes x_i ; $i = 1, 2, \dots, m$ of 1.3 (Guest, 1958). Also based on a classical regression model, Arndt et al. (2006) associated the problem of Gaussian node determination to the collection of survey data. They concluded that use of Gaussian nodes expedited previous costly and time-consuming data collection processes.

Among the four positions marked with the help of tabled values, shoot position A was closest to the trunk, position B followed by position C was next, and shoot position D was the location of the sampled shoot furthest from the trunk along the cordon. Clusters were then chosen randomly from the systematically selected shoots (primary sampling units) on a given cordon. Overall, 12 rows and 48 vines were selected from each block. Midseason, two clusters had been sampled per vine. In this way, 192 clusters had been sampled from each block. Later, at harvest, the remaining clusters were collected in boxes left beneath each previously sampled vine.

Besides determining each sampled cluster's weight, its rachis' main axis had been measured in three ways. As sketched in Figure 2, the narrowest rachis diameter, D_{\min} , between the point of attachment to the shoot and the small bump where a wing (if present) would attach was determined. Because the top part of the rachis is hourglass-shaped, this narrowest point, P, is easy to locate and thus,

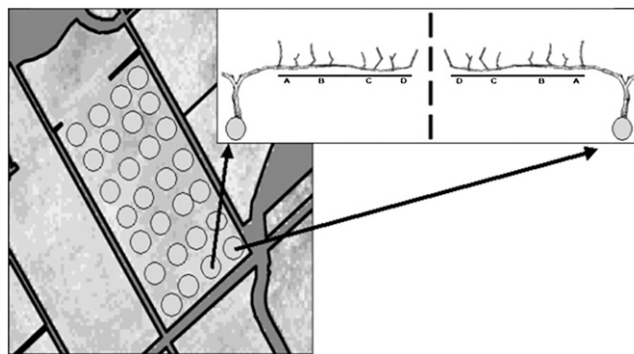


Fig. 1. Representation of one of five sampled vineyard blocks with vines marked as circles.

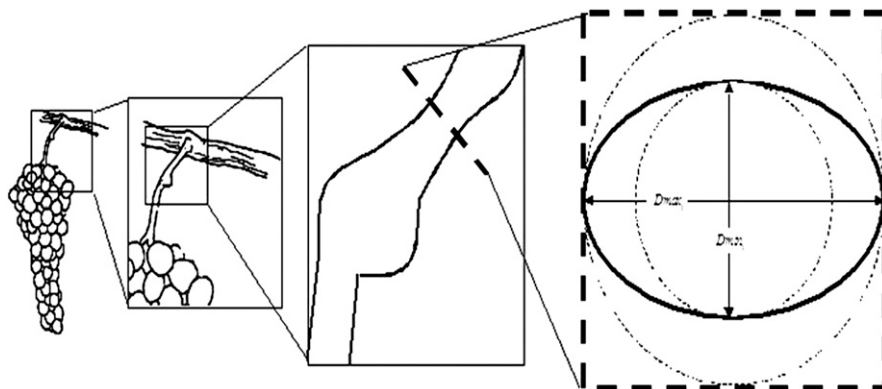


Fig. 2. Enlarged and idealized cross-section around the narrowest point on a rachis' top segment. Two measured diameters, D_{\min} and D_{\max} , are shown within the smallest square that encloses the cross section.

in this way, 192 clusters' rachis diameters were measured with a digital caliper. The rachis was then cut through P, and once all of the clusters' side branches and berries were removed so that just the rachis stalk remained, this stalk was weighed. Then, to straighten it as much as was feasible, the length of each rachis stalk was measured after pushing it into a plastic corner with measuring tape attached.

To obtain supplemental information, during the period that extended from 1 Sept. 2004 through the last date of harvest, 4 Oct. 2004 (Stages 35 and 38 according to Eichhorn and Lonertz classification), 96 clusters from CS1 and 96 clusters from CS2 were sampled. As before, vines and clusters were selected by using a random-systematic-random approach, and wing presence was recorded. On 14, 17, 21, and 24 Sept. from each of the two blocks, 24 clusters had been sampled randomly. Four berries were then removed from each selected cluster without removing the cluster from the shoot to which it was attached.

The particular berries selected had been chosen by a process designed to avoid statistical bias resulting from the sampling of berries grown near the bottom of the cluster (Tarter and Keuter, 2005). Once they had been removed from the cluster, these berries were weighed and °Brix was determined (from the juice obtained by crushing the berries together). Later, at harvest, the remaining berries of each previously sampled cluster were removed, weighed, crushed, and their combined °Brix determined.

Statistical methods. As illustrated by Hall et al. (2008), a simple transformation (Fisher's z) can improve the accuracy of vineyard-related studies that make use of the conventional Pearson correlation coefficient, r . To supplement Fisher's z -transformed values, Student's t , and analysis of variance (ANOVA) analytical methods, box and whisker plots (Tukey, 1977) displayed differences between wing-bearing and non-wing-bearing clusters sampled at midseason. Box and whisker plots supplemented analytical methods by displaying first and third quartiles and, unlike ANOVA techniques that assessed differences between population means, they estimated population medians.

Two types of correlation were computed. Type 1 correlations, r_{1all} and r_{2all} , were computed using measurements of all clusters, whereas Type 2 correlations, r_{1wing} and r_{2wing} , were computed using measurements of clusters that had wings, exclusively. Correlations r_{1all} and r_{1wing} assessed the degree of association of a vine's cluster's weight at midseason (wing included when present) multiplied by the number of clusters on the vine on which the cluster had grown and this same vine's yield at harvest. Correlations r_{2all} and r_{2wing} assessed the degree of association of a cluster's weight with this same vine's average, i.e., per cluster, yield at harvest. The pair of correlations r_{1all} and r_{1wing} differs from the pair of correlations r_{2all} and r_{2wing} in terms of the emphases placed on cluster count variation from vine

to vine. Specifically, *r1all* and *r1wing* tend to place emphasis on the variability of highly productive vines. Conversely, when per-cluster yield determinations involve low-valued counts, variation in these counts can affect *r2all* and *r2wing* substantially. Also, as discussed subsequently, because correlations *r1all* and *r1wing* pertain to per-vine yield, not per-cluster yield, the effects that nonproductive vines (such as those damaged by animal predation) can have on a sampled row's yield can be assessed easily. Null hypotheses of zero correlation-type difference were tested with Student's *t* tests, 4 df, the number of independently sampled blocks, here five, minus one.

The following two study characteristics are noteworthy: 1) when differences between two correlations were assessed, both correlations were calculated using measurements gathered concurrently. Hence, for the purposes of our study, data collection date differences had little effect on findings; and 2) all correlations pertained to the association between the weight of a single cluster removed from a vine at midseason and the harvest weight of all the vine's remaining clusters enumerated as follows. At midseason, the following mean numbers of clusters were counted: on CS1, 26.77; on CS2, 14.35; on ME1, 25.97; on ME2, 24.39; and on CH, 16.22.

When the population correlation coefficient is positive, the Pearson sample correlation coefficient *r*'s distribution is negatively skewed. Partly because of this skewness, Fisher introduced the transformation $z = (1/2)\log_e[(1+r)/(1-r)]$ and Fisher showed that *z* was approximately normally distributed (Schork and Remington, 2000). As a result of skewness, therefore, Pearson coefficients were transformed to their Fisher's *z* counterparts. Fisher's *z* helps improve study accuracy; because the quantity obtained by computing Fisher's *z* has a nearly normal sampling distribution, Fisher's *z*-transformed correlations will tend to have probability density curves that are more bell-shaped than the density curves that describe the distributions of their pre-transformed counterparts (Hoyle, 1973; Tarter, 2008).

Given that two variates, z_1 and z_2 , have a joint bivariate normal distribution, and that the correlation of z_1 and z_2 does not equal one, any difference between these variates, irrespective of their mutual dependence, must have a normal distribution and, therefore, the quantity $(z_1 - z_2)$ satisfies a condition required for Student's *t* test validity. This follows from Theorem 2.4.1 of Anderson (1984) when the first row vector of a matrix that Anderson designates as *C* assumes the value 1 as its first vector element and the value -1 as its second vector element. That $(z_1 - z_2)$ replicates (one replicate for each independently sampled block) that we computed could, for all practical purposes, be taken to be independent and identically distributed normal variates was checked using a graphical procedure illustrated, in the context of correlation coefficients, by Hill (1969).

The approach that we used to gather sampled measurements was designed so that

only two clusters were removed from any single vine sampled at midseason. The primary reason for this strategy was to reduce the possible impact that midseason cluster

removal from a vine had on this vine's yield. However, a second benefit of this approach was that substantial care could be taken to minimize any effect that a sampled cluster's

Table 1. Values used to assess wing presence's influence on yield-related correlations.²

Block	<i>r1all</i>	<i>r1nowing</i>	<i>r1wing</i>	<i>r2all</i>	<i>r2nowing</i>	<i>r2wing</i>	<i>nnowing</i>	<i>nwing</i>
CH	0.699	0.667	0.771	0.268	0.226	0.336	154	37
CS1	0.391	0.397	0.434	0.224	0.242	0.219	130	61
ME1	0.335	0.392	0.427	0.234	0.266	0.319	92	100
ME2	0.635	0.609	0.759	0.509	0.480	0.642	114	78
CS2	0.591	0.556	0.624	0.240	0.162	0.342	87	104

²Correlations *r1all* and *r2all* were computed using data about all clusters. *r1nowing* and *r2nowing* were computed using measurements of clusters that never had wings. *r1wing* and *r2wing* were computed using data restricted to clusters that had wings. All *r1* correlations were based on the dependent variate, total vine yield. All *r2* correlations were based on the dependent variate, per-cluster vine yield. *nnowing* specifies the number of the block's sampled clusters that never had wings. *nwing* specifies the number of the block's sampled clusters that had wings. (From the CH, CS1, and CS2 blocks, one cluster had been so damaged in transit that its weight could not be calculated.)

Table 2. Univariate analysis of variance assessed the factor, wing presence/absence (WINGPRES), in terms of this factor's relationship to the dependent variate, cluster weight (after a wing, when present, had been removed).²

Source	Type III sum of squares	df	Mean square	F	Significance
Corrected model	432,860.299 ^y	9	48,095.589	35.593	0.000
Intercept	3,299,046.508	1	3,299,046.508	2,441.431	0.000
WINGPRES	12,328.800	1	12,328.800	9.124	0.003
BLOCK	386,727.844	4	96,681.961	71.549	0.000
WINGPRES*BLOCK	13,503.368	4	3,375.842	2.498	0.041
Error	1,278,306.620	946	1,351.275		
Total	5,485,651.546	956			
Corrected total	1,711,166.919	955			

²The value of the dependent variate is computed after subtracting the weight of wings in cases where wings were present. The effect of prior wing presence is highly significant (significance level is 0.003).

^y $R^2 = 0.253$ (adjusted $R^2 = 0.246$).

Table 3. Univariate analysis of variance assessed WINGPRES in terms of this factor's relationship to the dependent variate, cluster weight (before a wing, when present, had been removed).²

Source	Type III sum of squares	df	Mean square	F	Significance
Corrected model	575,181.893 ^y	9	63,909.099	47.685	0.000
Intercept	3,894,048.992	1	3,894,048.992	2,905.507	0.000
WINGPRES	89,506.521	1	89,506.521	66.784	0.000
BLOCK	435,288.520	4	108,822.130	81.197	0.000
WINGPRES*BLOCK	28,099.803	4	7,024.951	5.242	0.000
Error	1,258,476.308	939	1,340.230		
Total	6,110,940.831	949			
Corrected total	1,833,658.201	948			

²The value of the dependent variate is computed with the weight of wings in cases where wings were present.

^y $R^2 = 0.314$ (adjusted $R^2 = 0.307$).

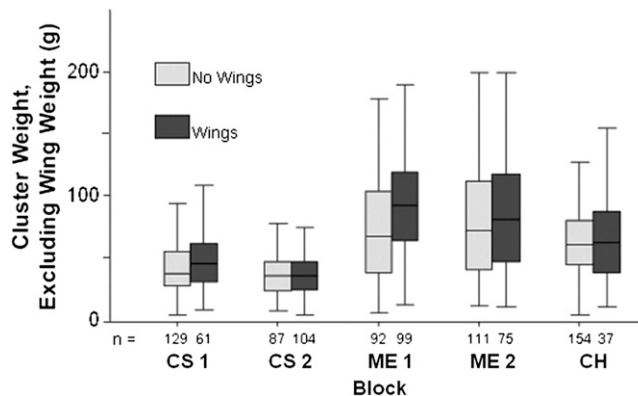


Fig. 3. Cluster weight box and whisker plots with midseason cluster weight computed after subtracting the weight of wings (in cases where wings were present) for the five sampled blocks. Sample sizes are marked below the X-axis. Box edges denote the first and third quartiles.

position along the cordon on which it grew might have on both the cluster's wing-related structure and its weight. To assess the relationship between shoot location, A, B, C, and D as specified previously, and the odds that a cluster will bear a wing and, in addition, to study the relationship between wing propensity and row location within a vineyard, χ^2 -based cross-tabulation approaches were applied as described by Cochran (1954).

Results

Correlation differences. Correlation coefficients determined solely from the subset of sampled clusters on which a wing was

present were significantly larger ($P = 0.014$) than coefficients determined from all sampled clusters. In a similar way, correlations between midseason cluster weight determinations and the sampled vine's average per cluster yield at harvest also differed significantly ($P = 0.023$). Table 1 lists the Pearson correlations from which the Student's t tests had been calculated. In only one instance, out of a possible 10, was a correlation based on data pertaining to clusters that bore wings calculated to be smaller than the adjoining table entry determined irrespective of wing presence or absence. This difference was, however, small, 0.224 for all clusters and 0.219 for clusters with wings.

Weight difference study. As listed in ANOVA in Table 2, despite wing removal before their being weighed, clusters that had once had wings were found to be significantly heavier than clusters grown without wings; a two-factor ANOVA disclosed a highly significant wing presence/absence (WINGPRES) effect ($P = 0.003$). ANOVA in Table 3 is the counterpart to ANOVA in Table 2 that pertains to cluster weights determined before wing removal, which, in light of findings listed in Table 1, not surprisingly also discloses a significant WINGPRES effect ($P = 0.000$).

Box and whisker plots in Figures 3 and 4 indicated that the wing presence with cluster weight relationship is less pronounced for the 'Cabernet Sauvignon' blocks than it is for the 'Merlot' and 'Chardonnay' blocks. With this exception, even when weights were measured after the removal of wings from wing-bearing clusters, wing-bearing clusters were found to be significantly heavier than their wingless counterparts. This finding, which is shown clearly by the heights of the plots' central (i.e., median) bars in Figures 3 and 4, prompted further study of late-season measurements that had been gathered from the CS2 and CS1 blocks. Figure 5 shows cluster weight box and whisker plots for the CS1 and CS2 sampled blocks. Median bar differences between plot pair members are notable. For instance, regarding the plot of CS1 cluster weights (determined after wing removal), the difference between the two median bars constructed using midseason data are nearly

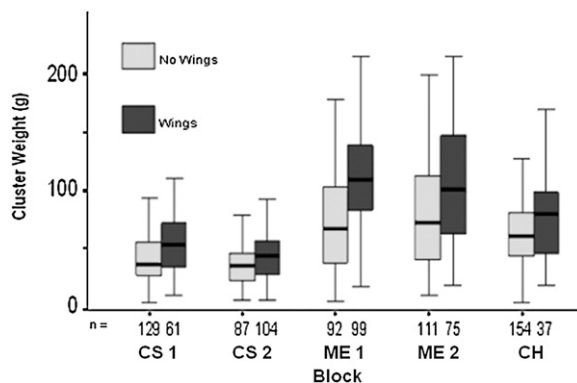


Fig. 4. Cluster weight box and whisker plots for the five sampled blocks. In cases where wings were present, midseason cluster weight is computed with the weight of wings included. Sample sizes are marked below the X-axis. Box edges denote the first and third quartiles.

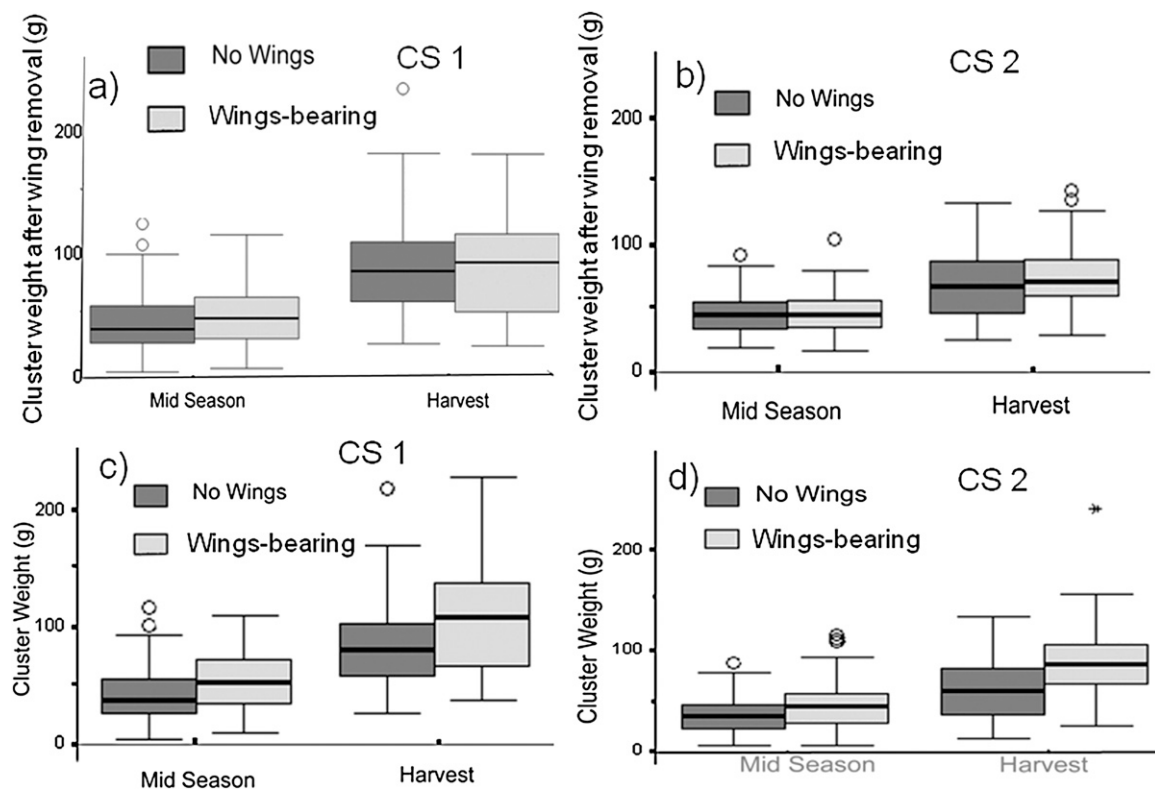


Fig. 5. Cluster weight box and whisker plots for the Dm-CS and Tw-CS sampled blocks. As shown on the upper plots (A–B), cluster weight was computed after subtracting the weight of wings (when wings were present). The lower plots (C–D) were computed before subtracting the weight of wings. Any circle outside the box has a value either smaller than the first quartile or, alternatively, larger than the third quartile.

the same as the difference between the two median bars determined from measurements taken at harvest. It was also found that among the four pairs of plots based on weights determined after wing removal, no rightmost pair member differs substantially from its leftmost plot counterpart (although leftmost and rightmost plots were constructed using data from independent samples).

For the CS1 and CS2 sampled blocks, as shown in Figure 6, based on measurements of four berries that had been removed from each sampled cluster (samples were collected so that none of the four berries was grown on a wing, if present), estimated average berry weight box and whisker plots were constructed. The only notable difference between a plot based on data from wing-bearing clusters and a plot based on data from non-wing-bearing clusters was (in comparison with its non-wing-bearing counterpart) the substantially greater interquartile range of the CS2 plot constructed using measurements of wing-bearing clusters. Despite this observed difference, all four of the plots' median bars assume nearly identical values.

Based on measurements gathered as sketched in Figure 2, 10 box and whisker plots are shown in Figure 7. All five WINGPRES box and whisker plot Dmin value median bars were found to assume higher values than the values attained by their non-WINGPRES box and whisker plot median bar counterparts. As illustrated by the box and whisker plots shown in Figure 8, with the exception of the median bars computed from CS1-gathered measurements, rachis weight and rachis length measurement-related median bars were found to assume higher values than the values attained by their non-WINGPRES box and whisker plot median bar counterparts.

Wing presence and late-season °Brix. Once wings had been separated from CS1 and CS2 clusters that had been selected on 4 Sept., after crushing all its berries, the °Brix of each wing had been measured. This value was then subtracted from the counterpart measurement of cluster °Brix (wing berries excluded) also taken on 4 Sept. The box and whisker plots of Figure 9 show estimates of this difference variate's (DBRIX) distribution that were obtained from CS1 and CS2 measurements, respectively. DBRIX was found to be distributed with a nearly zero estimated median value. When estimated from the CS1 and CS2 berries that we selected on 4 Sept., and with one important proviso, the °Brix variate's distribution appears to depend minimally on whether a berry is grown on a wing.

The distinction between mean-based and median-based analyses underlies the proviso referred to in the preceding paragraph. The upper whiskers shown in Figure 9 are longer than their lower counterpart whiskers. This is a strong indication of probability density function (pdf) asymmetry, which, when the right curve tail is heavier than the left curve tail, tends to place means, i.e., arithmetic averages, to the right of their counterpart medians. To compensate for distribution asymmetry, a pro-

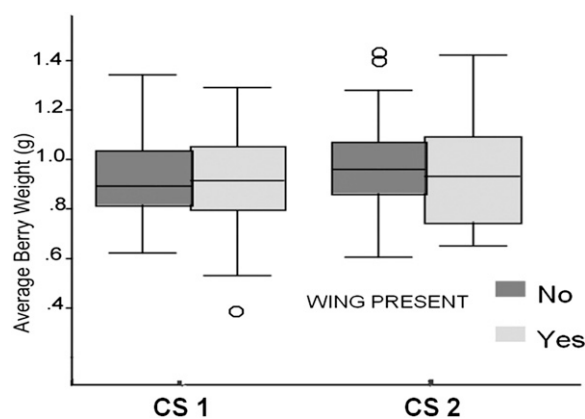


Fig. 6. Average berry weight box and whisker plots based on data sampled from the Tw-CS and Dm-CS blocks for clusters with or without wings. Any circle outside the box has a value either smaller than the first quartile or, alternatively, larger than the third quartile.

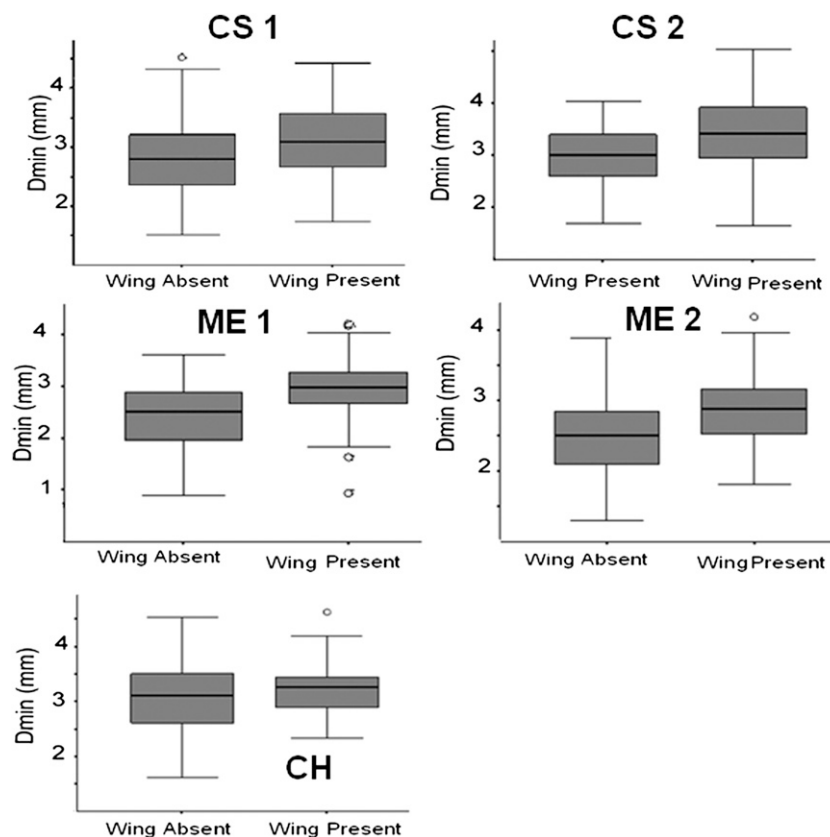


Fig. 7. Rachis minimum diameter (Dmin) box and whisker plots in clusters with or without wings for the five blocks. Any circle outside the box has a value either smaller than the first quartile or, alternatively, larger than the third quartile.

cedure whose findings are summarized in ANOVA in Table 4 was implemented. The dependent variable whose distribution is assessed with the help of Table 4 is the arithmetic average of the two °Brix measurements. The first measurement was based on four crushed berries sampled a few days before cluster removal. After crushing all of the cluster's remaining berries, from each of the CS2 and CS1 clusters that had been selected on 4 Sept., a second °Brix determination had been made (after the cluster's wing, when present, had been separated). The two

main effects listed in Table 4 were, WINGPRES, i.e., whether a measured cluster had a wing and, BLOCK, the factor that distinguished the CS2 block from the CS1 block.

Cluster's position along the cordon and row's position within the vineyard. To assess the relationship between a shoot's location along the cordon to which it was attached and whether the shoot bears clusters with wings, cross-tabulations were computed, as listed in Table 5. Among the sampled shoots, the shoot denoted as the A of Table 5 grew closest to the sampled vine's trunk, whereas

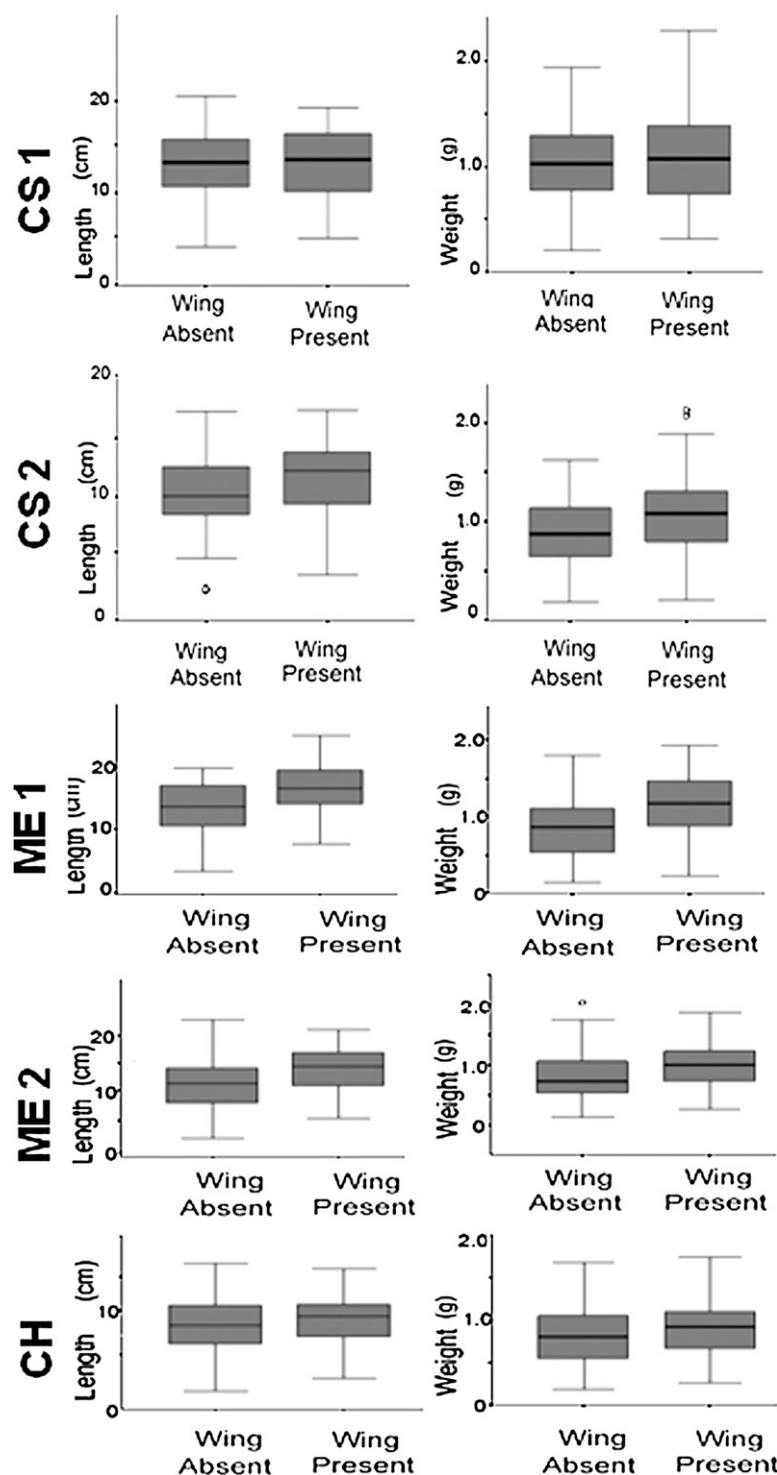


Fig. 8. Stretched rachis length (leftmost plots) and rachis weight (rightmost plots) box and whisker plots in clusters with or without wings for the five blocks. Any circle outside the box has a value either smaller than the first quartile or, alternatively, larger than the third quartile.

the shoot sampled from cordon location D was furthest from the trunk. χ^2 statistics were calculated to assume the small values, 1.53, 1.23, 3.17, 3.02, and 0.37. Given the df available, for any of these values to have been significant at a 0.01 P value, the calculated χ^2 would have had to be 11.34 or higher.

Cross-tabulations, similar to those listed in Table 5 were calculated to assess the relationship between, on the one hand, a sampled row's location within the CS1, CS2, ME1,

ME2, and CH block and, on the other hand, wing presence or absence. The differences between the χ^2 calculated, 36.85, 17.56, 40.24, 8.03, and 22.06, were notable. For any of these values to be significant at a 0.01 P value, the calculated χ^2 would have to be 24.72.

Discussion

The finding that a °Brix variate's distribution shows little dependence on whether

a berry is grown on a wing has two practical implications. The first is relevant whenever a seasonal berry sample is gathered for the purpose of plotting ripening curves and/or assessing maturation in the field. Especially in cases of high cluster compactness, such a sample might rely on detachment of wing berries, which are notoriously more easily reachable than those inserted along the main axis. The second implication is relevant whenever within-cluster thinning is applied (like it still is in some premium wine areas). At least for the varieties we studied, selective wing removal under the assumption that these wings might carry berries that are somewhat less uniform (in the sense of retarded ripening) was found to have little justification.

The larger values displayed in Table 1 for r1all, r1nowing, and r1wing determinations, when these values are compared with counterpart values displayed for r2all, r2nowing, and r2wing, are noteworthy in terms of the emphases placed on cluster count variation from vine to vine. The two correlations, r1all and r1wing, emphasize the variability of highly productive vines, whereas low-valued count variation will affect r2all and r2wing substantially. Obviously within a specific growing season, for vineyard-practice reasons, it is preferable to pay more heed to high-productivity vines than to their lesser yield counterparts.

A random-systematic-random sampling method had helped to account for non-rectangular block geometry (Tarter and Keuter, 2008). While implementing procedures for measuring vines that had been chosen randomly from each sampled row, it required little additional effort to count the number of nonproductive vines within the selected row. Because only productive vines were sampled at midseason, this count had the potential to provide an estimate of the sampled row's productivity. At harvest, our investigations could take advantage of each previously sampled vine's remaining clusters being collected in boxes left beneath the vine. Hence, we had the luxury of being able to correlate vine-specific, not row-specific, information. As a consequence, we could bypass the productive vine count issue, an issue that if ignored can bias findings substantially. Specifically, were a per-vine yield estimate multiplied by the total number of vines within the sampled row, not the total number of productive vines, this would inflate estimates of the row's true yield.

There is a plausible statistical reason for the larger values displayed in Table 1 for r1all, r1nowing, and r1wing determinations when these values are compared with counterpart values displayed for r2all, r2nowing, and r2wing. These differences can be explained in terms of combinations of variates that—individually—are distributed with a normal probability density function. Given that two variates have a joint bivariate normal distribution, although sums and differences of these variates will always have a univariate normal pdf (provided that their correlation's absolute value is less than one), even when these variates are correlated, have means equal to zero,

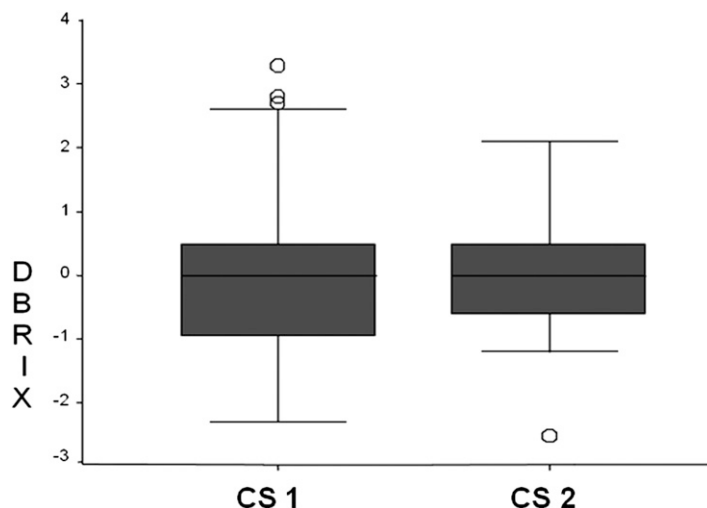


Fig. 9. Difference in °Brix (°Brix of remainder of cluster after wing removal, minus Brix of wing); box and whisker plots constructed using harvest data from Tw-CS and Dm-CS sampled blocks. Any circle outside the box has a value either smaller than the first quartile or, alternatively, larger than the third quartile.

Table 4. Univariate analysis of variance whose main effects were the factor, WINGPRES, i.e., whether a measured cluster had a wing, and a factor whose levels distinguished the CS2 block from the CS1 block (BLOCK).^z

Source	Type III sum of squares	df	Mean square	F	Significance
Corrected model	127.763 ^y	3	42.588	13.495	0.000
Intercept	112,012.874	1	112,012.874	35,495.473	0.000
WINGPRES	10.634	1	10.634	3.370	0.068
BLOCK	83.251	1	83.251	26.381	0.000
WINGPRES*BLOCK	0.001	1	0.001	0.000	0.983
Error	593.271	188	3.156		
Total	128,446.52	192			
Corrected total	721.033	191			

^yDependent variable was AVEBRIX.

^z $R^2 = 0.177$ (adjusted $R^2 = 0.164$).

and have SDs equal to one, i.e., both variates are standard normal distributed, their quotient will have a 1 df Student's *t* probability density function. This particular Student's *t* probability density has curve tails that differ from normal curve tails substantially.

As illustrated by Stigler (1974), the 1 df Student's *t* distribution has many unusual characteristics. For example, it has an infinite population SD. Although none of the quantities whose correlations are estimated by *r*_{2all} and *r*_{2wing} can have an infinite population SD, among possible reasons why *r*_{1all}, *r*_{1nowing}, and *r*_{1wing} attained values larger than those attained by their *r*_{2all}, *r*_{2nowing}, and *r*_{2wing} counterparts is the tendency for quotients whose denominators vary substantially to accentuate statistical noise caused by a large-valued population SD.

Regarding the 4 df Student's *t* distribution on which the finding, "Correlation coefficients determined solely from the subset of sampled clusters on which a wing was present were significantly larger (p-value 0.014) than coefficients determined from all sampled clusters," was determined, a checking procedure was implemented. Ordinarily, a Student's *t* test is based, in part, on the selection of a sample SD (here *s* = 0.109) as an estimator of a population SD (here the unknown value, σ). As an alternative, Table 10C.1 of Sarhan and Greenberg

(1962) was used to compute a best linear unbiased (BLU) estimator of σ (here calculated to equal 0.114). The similarity of the two values, 0.109 and 0.114, provided a validity control. (Because they are not determined from squared sample values, BLU estimators of scale tend to be less swayed by sample extremes than statistics *s*.)

Statistical as well as vineyard-related reasons motivated averaging the measurements of °Brix taken on distinct dates on which the box and whisker plots shown in Figure 9 were based. There is no reason to believe that a °Brix-related measurement's distribution has a symmetric pdf. Usually, however, the designation, symmetric and bell-shaped, will be more appropriate when applied to an average of any two component measurements, be they correlated or uncorrelated, than it will be when applied to this average's components individually. Hence, it is advantageous to average °Brix measurements. In addition to statistical considerations, there is a practical reason for gathering °Brix-related measurements from a single cluster that is sampled at different dates. It is obviously impossible to sample all clusters concurrently. However, randomly sampling a cluster's berries at four different dates and then averaging measurements can reduce the confounding that would otherwise be attributable to late-season day-to-day °Brix changes.

Table 5. Cross-tabulations, i.e., the factor, whether a cluster had a wing, tabulated with respect to the sampled shoot's location on a cordon.^z

Wing present ^y	Position on the cordon: A, B, C, and D.			
	A	B	C	D
No	22, 34, 20, 30, 40	25, 31, 25, 32, 38	19, 35, 27, 28, 38	21, 31, 20, 24, 39
Yes	26, 14, 28, 18, 8	23, 17, 23, 16, 10	29, 13, 21, 20, 10	26, 17, 28, 24, 9
Total	48, 48, 48, 48, 48	48, 48, 48, 48, 48	47, 48, 48, 48, 48	48, 48, 48, 48, 48
Total	70	73	66	78

^yThe symbol A denotes the location closest to the trunk, through D, furthest from the trunk. Notable are the small χ^2 calculated, 1.53, 1.23, 3.17, 3.02, and 0.37. For any of these values to have been significant at a 0.01 *P* value, the calculated χ^2 would have had to be 11.34. (Table entries and χ^2 statistics, 1.53, 1.23, 3.17, 3.02, and 0.37, are listed in the following order, first CS1, ME1, ME2 and, finally, CH.)

Literature Cited

- Anderson, T.W. 1984. An introduction to multivariate statistical analysis. John Wiley & Sons, New York, NY.
- Arndt, C., J. Kozlitina, and P.V. Preckel. 2006. Efficient survey of households via Gaussian quadrature. *J.R. Stat. Soc.* 55:355–364.
- Carmo Vasconcelos, M., M. Greven, C.S. Winefield, M.C.T. Trought, and V. Raw. 2009. The flowering process of *Vitis vinifera*: A review. *Amer. J. Enol. Viticult.* 60:411–434.
- Clingeffer, P.R., M. Kristic, and K. Sommer. 2000. Production efficiency and relationships among crop load, fruit composition and wine quality. Proceedings 50th Anniversary of the American Society of Enology and Viticulture, Annual Meeting of the American Society for Viticulture and Enology: Seattle, WA, pp. 318–322.
- Cochran, W.G. 1954. Methods for strengthening the common χ^2 tests. *Biometrics* 10:417–451.
- Coombe, B.G. 1995. Growth stages of the grapevine: Adoption of a system for identifying grapevine growth stages. *Aust. J. Grape Wine Res.* 2:104–110.
- Dunn, G.M. and S.R. Martin. 2007. A functional association in *Vitis vinifera* L. cv. Cabernet Sauvignon between the extent of primary branching and the number of flowers formed per inflorescence. *Aust. J. Grape Wine Res.* 13:95–100.
- Guest, P.G. 1958. The spacing of observations in polynomial regression. *Ann. Math. Stat.* 29: 294–299.
- Hall, A., J.P. Louis, and D.W. Lamb. 2008. Low-resolution remotely sensed images of wine grape vineyards map spatial variability in planimetric canopy area instead of leaf area index. *Aust. J. Grape Wine Res.* 14:9–17.
- Hill, A. 1969. On looking at large correlation matrices. *Biometrika* 56:249–253.
- Hoel, P.G. 1958. Efficiency problems in polynomial estimation. *Ann. Math. Stat.* 29:1134–1145.
- Hoel, P.G. and A. Levine. 1964. Optimal spacing and weighting in polynomial prediction. *Ann. Math. Stat.* 35:1553–1560.
- Hoyle, M.H. 1973. Transformations: An introduction and bibliography. *Int. Stat. Rev.* 41:203–223.
- Pratt, C. 1971. Reproductive anatomy in cultivated grapes—A review. *Amer. J. Enol. Viticult.* 22: 92–109.
- Sarhan, A.E. and B.G. Greenberg. 1962. Contributions to order statistics. Wiley, New York, NY.
- Schork, M.A. and R.D. Remington. 2000. Statistics with applications to the biological and health sciences. 3rd Ed. Prentice Hall, New York, NY.
- Stigler, S.M. 1974. Studies in the history of probability and statistics, XXXIII—Cauchy and the witch of Agnesi: An historical note on the Cauchy distribution. *Biometrika* 61: 375–379.
- Stroud, A.H. and D. Secrest. 1966. Gaussian quadrature formulas. Prentice Hall, New York, NY.
- Tarter, M.E. 2008. Data transformations, p. 249–254. In: Boslaugh, S. (ed.). *Enc. epidemiology*. Sage Publications, Thousand Oaks, CA.
- Tarter, M.E. and S.E. Keuter. 2005. The effect of rachis position on size and maturity of Cabernet Sauvignon berries. *Amer. J. Enol. Viticult.* 56: 86–89.
- Tarter, M.E. and S.E. Keuter. 2008. Shoot-based sampling of *Vitis vinifera* clusters. *Amer. J. Enol. Viticult.* 59:55–60.
- Tukey, J.W. 1977. *Exploratory data analysis*. Addison-Wesley, Reading, MA.