Easy and accurate estimation of grapevine leaf area with simple mathematical models

C. LOPES and P. A. PINTO

Universidade Técnica de Lisboa, Instituto Superior de Agronomia, Lisboa, Portugal

Summary

Empirical models for non-destructive estimation of primary and lateral leaf area are presented for cv. Aragonez, syn. Tempranillo. Primary leaf area of shoots is estimated by a model using a calculated variable obtained from the average of the largest and smallest primary leaf area multiplied by the number of primary leaves. For lateral leaf area estimation another model is presented which uses the same type of calculated variable plus the area of the largest leaf. Both models explain a high proportion of the leaf area variability and have a good predicting capability. Validation with 4 independent data sets periodically sampled during several seasons from different varieties shows that these models can predict primary and lateral leaf area independently of variety, year and growth stage, using a non-destructive, low cost, simple and accurate method. The models represent a powerful tool for grapevine research, for consultants and advanced growers allowing the evaluation of leaf area more frequently.

Key words: grapevine, leaf area, statistical model, Tempranillo, validation.

Introduction

In viticulture the leaf area, its density and distribution within the canopy are fundamental parameters for characterizing the light microclimate and to understand the responses of vines to environment, training systems and canopy managements. The monitoring of grapevine leaf area is, however, not an easy task.

A variety of techniques have been proposed to estimate grapevine leaf area. Indirect, non-destructive estimation of leaf area can be done, e.g. by measuring light extinction through the canopy (Grantz and Williams 1993; Sommer and Lang 1994; Oliveira and Santos 1995; Ollat et al. 1998; Patakas and Noitsakis 1999), by empirical models where leaf area development is defined as a function of temperature (Schultz 1992; Bindi et al. 1997), or by remote sensed imagery (Dobrowski et al. 2002). These methods are rapid but very expensive devices are necessary that need frequent calibration and specific sampling protocols (Ollat et al. 1998) and, for very dense canopies, they often underestimate leaf area (Cohen et al. 2000). Another disadvantage of these methods is the fact that they are unable to distinguish between primary and lateral leaf area, which is a strong handicap in viticultural studies.

Direct methods to determine leaf area are based on measurements of leaf and/or shoot samples, and are either destructive or non-destructive. Destructive methods are based on the removal and measurement of leaf area in the lab by planimeters, determination of the area to weight ratio (Sepulveda and Kliewer 1983) and area meters. These methods are easy and accurate (Sommer and Lang 1994) but are laborious, time-consuming and will reduce the photosynthetic leaf area.

Direct non-destructive methods consist of in situ leaf area measurement with portable devices, very expensive and difficult to use in the field. Other direct non-destructive methods are based on empirical relationships between leaf area and other parameters that can easily be measured directly on the canopy. To estimate the area of single vine leaves several statistic models have been proposed, e.g. based on measurement of the two lateral leaf veins (Carbonneau 1976 a; Lopes and Pinto 2000), the length of the primary vein (Schultz 1992) or the length and/or the maximal leaf width (Smith and Kliewer 1984; Elsner and Jubb 1988). These methods are simple and accurate but in order to estimate total leaf area one still has to measure all primary and lateral leaf veins, a very laborious procedure.

To overcome this problem Carbonneau (1976 b) proposed to measure one leaf sample in each group of 4 contiguous leaves without loosing accuracy. Also Barbagallo et al. (1996) proposed an empirical model to estimate primary leaf area per shoot based only on the measurement of 3 leaves: the largest leaf, the apical leaf and an intermediate leaf. These methods allow an important reduction of the number of leaves measured but they were developed for primary leaf area estimation only and the authors do not refer its applicability for lateral leaf area estimation. Yet, lateral leaf area can represent an important proportion of the total leaf area, especially at high vigour (Huglin and Schneider 1998).

To estimate total leaf area per shoot, Mabrouk and Carbonneau (1996) have proposed a simple model based on correlations between total leaf area and the length of the primary and lateral shoots. The simplicity of this model makes it appropriate for common use since no special equipment is needed. Shoot length, however, is not always closely correlated with leaf area, especially for primary shoots (Lopes and Pinto 2000; Tregoat et al. 2001).

Lopes and Pinto (2000) proposed another empirical model for the estimation of primary leaf area per shoot, which includes 4 variables: shoot length, number of primary leaves and area of the largest and smallest leaves. For lateral leaf area they proposed a similar model that considers each lat-
eral shoot as a composed leaf. In spite of their good accuracy these models were developed from a set of shoot measurements at the end of the growing season, therefore lacking information on earlier growth stages.

The aim of this paper is to further develop the approach of Lopes and Pinto (2000) to obtain a simple and accurate model for easy and rapid estimation of primary and lateral leaf area independent of the vine growth stage. A second objective is to validate the obtained models with independent data from different varieties, years and locations.

**Material and Methods**

**Field conditions and plant material:** Shoots were collected in 1997 in a commercial vineyard located at Estremoz, southern Portugal, within the Alentejo wine region (38° 46’ N; 7° 29’ W). The 8-year-old vines, cv. Aragonez (syn. Tempranillo), grafted to 1103 Paulsen were spaced 2.7 m between and 1.1 m in rows and trained on a bilateral Royat cordon system using vertical shoot positioning with a pair of movable wires. The trunk height was 0.6 m, the vines were spur-pruned (8 spurs with 2 buds per vine) and the shoots were trimmed twice (May and June) at about 1.0 m height.

The soil is derived from schist with a variable depth (1.0-1.5 m). Soil horizons present a variable texture (clay fraction ranges from 20 to 34%), gravel and rock fragments on the surface.

**Leaf area measurements:** Periodically, from April 4 (3 weeks after bud burst) to the end of August (6 sampling dates), a sample of 30 fruiting shoots (one shoot per vine) was randomly collected from the outside and inside of the canopy, inserted in plastic bags and transported to the lab. From each shoot, primary and lateral leaves were separated and numbered according to node insertion and the leaf area was measured by a leaf area meter (Delta-T area meter, Delta-T Devices, England). When shoots had a primary leaf arising from a base bud this leaf was excluded as, in general, it was too small and had a very abnormal shape. Young leaves with a primary vein length <3 cm were also excluded. Primary shoot length (SLp) was measured from the base to the node where the last counted leaf was inserted.

Data from the 180 shoots measured were separated into the categories ‘primary’ and ‘lateral’. For each primary shoot the following variables were computed: sum of primary leaf area (LA1), number of primary leaves (NL1), area of the largest primary leaf (L1) - the highest primary single leaf area, area of the smallest primary leaf (S1) - the lowest single primary leaf area. From these variables two new variables were calculated: the mean primary leaf area: \( M_1 = (L_1 + S_1)/2 \) (Eq. 1); the mean primary leaf area per shoot: \( M_{LA1} = M_1 * NL_1 \) (Eq. 2).

For lateral leaf area the measurements started at the third sampling date and, since not all shoots carried laterals, only 107 primary shoots were used. All lateral leaves were grouped into one set of data from which the same type of variables reported for primary leaves were computed per shoot: sum of lateral leaf area (LAs), number of lateral leaves (NLs), area of the largest lateral leaf (Ls) and area of the smallest lateral leaf (Ss). A similar approach was used for calculated variables: the mean lateral leaf area: \( M_2 = (L_2 + S_2)/2 \) (Eq. 3); the mean lateral leaf area per shoot: \( M_{LA2} = M_2 * NL_2 \) (Eq. 4).

For model validation 4 sets of shoots were used; they were collected periodically in commercial vineyards from bud burst to harvest, from different red varieties in different seasons and wine regions: 130 shoots of cv. Aragonez from the same wine region and season (1997) but different locations (Évora, latitude 39° 04’ N); 80 shoots of cv. Cabernet-Sauvignon in 2002 from the Estremadura wine region (latitude 39° 04’ N); 168 shoots of cv. Touriga Nacional in 1994 and 230 shoots of cv. Jaen in 2001, both from the Dão wine region (latitude 40° 0’ N). Vines were more than 6 years old, trained vertically with shoot positioning and spur-pruned on a Royat bilateral cordon.

**Statistical analysis:** For each type of leaf area a multiple regression analysis between leaf area per shoot (dependent variable) and all the measured and calculated variables (independent variables) was performed using SAS® statistical software. Variable selection was done with a forward stepwise regression with 0.15 critical F statistic.

For validations, in order to evaluate the models goodness of fit to the observed data, the following deviance measures were used (Schaefhfer 1980): mean absolute error: \( MAE = \text{Σ} |y_i - \hat{y}_i|/n \) (Eq. 5); mean absolute percent error: \( MAPE = 100 * \text{Σ} |y_i - \hat{y}_i|/|y_i|/n \) (Eq. 6) where \( y_i \) represents observed values, \( \hat{y}_i \) simulated values and \( n \) the number of pairs.

The linear regression analysis of observed vs. predicted and the modeling efficiency (EF) were also used. For the regression analysis the observations were taken as Y-variates (Mayer and Butler 1993) and the F-test for slope = 1 and for intercept = 0 were performed using SAS®. The EF is a dimensionless statistic that relates model predictions to observed data and is defined as (Logague and Green 1991): \( EF = 1 - \Sigma (y_i - \hat{y}_i)^2/\Sigma (y_i - \bar{y})^2, \) (Eq. 7) where \( \bar{y} \) represents the mean of the observed values.

**Results**

**Growth of primary leaves:** The 1997 growth cycle started earlier than average with bud burst at the beginning of March, bloom at the end of April and veraison in the second week of July. Grapes were harvested at the end of August when sugar content had reached 24 °Brix. The growth of primary leaf area presented a normal pattern with an increase in spring and a plateau thereafter, followed by a small decrease during berry ripening. Lateral leaf area was highest at the end of May followed by a small decrease caused by a second pruning in mid-June, and a plateau thereafter. Shoot length was highest before the first pruning (beginning of May); it stabilized thereafter and so did the number of primary leaves which on average was 12 (Fig. 1).

For most primary shoots the size of primary leaves increased from node 1 to nodes 3-6 and then decreased in apical direction (Fig. 2). During the growth period, the largest primary leaf was between node 2 and node 9, most often
between nodes 3 and 6 (Fig. 3 A). Between bud burst and bloom most shoots showed the smallest measured leaf inserted at the apex; nevertheless, after bloom, the smallest leaves were also found at the basal part of the shoot or in other node positions (Fig. 3 B).

**Estimation of primary leaf area:** The correlation matrix between primary leaf area (LA1) and the 6 selected variables is shown in Tab. 1. The highest correlation coefficient was obtained with MLA1, followed by SL1, L1, M1 and NL1; the smallest correlation coefficient was obtained with S1.

To find an appropriate set of independent variables to predict the values of the primary leaf area, a forward stepwise regression analysis between LA1 (dependent variable) and 5 independent variables was performed. In order to avoid collinearity problems the calculated variable M1 was excluded as it is a linear combination of the two measured variables L1 and S1. The first variable entered into the model was MLA1, explaining a very high proportion of primary leaf area variability (R² = 0.97). No other variable met the 0.15 significance level for entry into the model which is represented by the following equation: LA1 = 49.1936 + 0.9958 * MLA1 (Eq. 8); R² = 0.97 (p < 0.001); n = 180; RMSE = 150.6 cm² where LA1 is the primary leaf area per shoot (cm²) and MLA1 the mean primary leaf area per shoot (cm²).

With the aim to detect significant interactions between the explanatory variables and the sampling date, a co-variance analysis was performed using the sampling date as a covariate. No significant interactions occurred.

The estimated values of the elected model fit very well with the actual leaf area (Fig. 4), however, residual plot shows that primary leaf area variation is dependent on the values of the predictor variable. The violation of the constant variance assumption indicated the need of a variable transformation. A logarithmic transformation of both sides of the
equation was applied to stabilize the variance (Eq. 9): LA1 = EXP[(0.0835 + 0.992 * ln(MLA1)] (Eq. 9); R2 = 0.99 (p < 0.001); n = 180; RMSE= 152.7 cm².

**Estimation of lateral leaf area:** For the estimation of lateral leaf area (LA2) the type of variables reported for primary leaves was used except for lateral shoot length. The correlation between LA2 and the 5 variables is shown in Tab. 2. As for primary shoots, the highest correlation coefficient was obtained with the calculated variable MLA2 (mean lateral leaf area per shoot) followed by NL2, L2 and M2. The variable S2 presented a very low and non-significant correlation coefficient with lateral leaf area.

<table>
<thead>
<tr>
<th>NL2</th>
<th>L2</th>
<th>S2</th>
<th>M2</th>
<th>MLA2</th>
<th>LA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.70</td>
<td>1</td>
<td>0.20</td>
<td>1</td>
<td>0.95</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Tab. 2

Likewise, MLA2 was the first variable selected by the stepwise regression to enter the model, explaining a very high proportion of lateral leaf area variability (partial R² = 0.97). In the second step variable L2 was chosen, but, in spite of its significance (prob. F = 0.04), its contribution to the explanation of lateral leaf area variability (partial R² = 0.001) is very low. No other variable met the 0.15 significance level needed to enter the model. The elected model is represented by the following equation: LA2 = 78.01726 – 1.16773 * L2 + 1.00848 * MLA2 (Eq. 10); R² = 0.97 (p<0.001); n = 107; RMSE= 97.9 cm² where LA2 is the lateral leaf area per shoot (cm²), L2 the area of the largest lateral leaf (cm²) and MLA2 the mean lateral leaf area per shoot (cm²).

As for primary shoots the co-variance analysis using the sampling date as covariate showed no significant interactions. A similar logarithmic transformation was also needed to stabilize the variance (Eq. 11). LA2 = EXP[(0.346 + 1.029 * ln(MLA2) – 0.125 * ln(L2)] (Eq. 11); Adj. R² = 0.98 (p<0.001); n = 107; RMSE = 93.2 cm².

**Model validation:** In order to test the potential applicability of the model to different locations and varieties, a validation was performed using independent datasets periodically sampled along the season. Visual observation showed a very good agreement between observed and estimated values for the 4 datasets and also a bias proportional to shoot leaf area (Fig. 5 A). The visual appraisal is corroborated by the statistic measures of validation (Tab. 3). The mean absolute percent error (MA%E) presented similar values between varieties (7 - 9 %). The linear regression between observed (dependent variable) and estimated (independent variable) shows a very high and significant R² for the 4 datasets, although the fitted lines presented some differences. While the fitted lines of cv. Aragonez (from Évora) and Jaen present an intercept not significantly different from the intercept of the reference line, the fitted line of cv. Touriga Nacional is significantly different from the reference line. The fitted line of cv. Cabernet-Sauvignon is not significantly different from the reference line.
zero, cv. Cabernet-Sauvignon and cv. Touriga Nacional fitted lines showed an intercept significantly different from zero. All the fitted lines presented a slope significantly >1. All the modeling efficiency values were higher than 0.90, the lowest value being the Cabernet-Sauvignon, the highest the Jaen data set. The combined set of data of the 4 varieties presented intermediate deviance measures, a fitted line with an intercept not significantly different from zero and a very high modeling efficiency.

In general the validation of the model for lateral leaf area presented similar goodness of fit as that reported for primary shoots (Fig. 5 B and Tab. 4). However, the MA%E present a higher variability between varieties and higher values than those obtained on the validation of primary leaf area. All linear regressions had a very high and significant R² and the intercepts of the fitted lines were not significantly different from zero with the exception of the Touriga Nacional data set. As for primary shoots, the slopes of all fitted lines were significantly different from 1; with the exception of the Aragonez data set, all slopes were <1. All modeling efficiencies were ≥ 0.90, the lowest value being the Touriga Nacional data set, the highest was the Aragonez data set collected at Évora. The combined set of data of the 4 varieties presents the lowest MA%E, the highest intercept being significantly different from zero, the highest RMSE and an intermediate modeling efficiency.

**Discussion**

The elected model for estimating primary leaf area is based only on one variable, the mean primary leaf area per shoot. The high importance of this variable can be derived from the measured variables from which it was calculated: a mean leaf area (largest + smallest/2) multiplied by the number of leaves. The importance of these three measured variables as predictors of the primary leaf area has already been reported. BARBAGALLO et al. (1996) included the area of the largest primary leaf in a model to predict primary leaf area. In a model to estimate primary leaf area proposed in a previous paper (LOPES and PINTO 2000), those variables were the first three variables selected by the stepwise regression to enter the model. In comparison, the new feature of our work is the composed variable that incorporates those three measured variables. The idea was to find a representative mean leaf area (largest leaf + smallest leaf/2) and then to multiply it by the number of leaves. The resulting calculated variable (MLA) is much closer correlated with the actual primary leaf area than any other measured variable alone, giving better accuracy to the model.

Shoot length is reported to be a measured variable very closely correlated with primary leaf area (SPARK and LARSEN 1966; MABROUK and CARBONNEAU 1996). In this work, however, shoot length was not selected as an explanatory vari-

<table>
<thead>
<tr>
<th>Data set</th>
<th>n</th>
<th>Dev. measures MAE</th>
<th>MA%E</th>
<th>R²</th>
<th>Linear regression Intercept¹</th>
<th>Slope²</th>
<th>RMSE</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aragonez</td>
<td>156</td>
<td>61.8</td>
<td>7.6</td>
<td>0.98***</td>
<td>0.64 ns</td>
<td>1.03*</td>
<td>99.5</td>
<td>0.97</td>
</tr>
<tr>
<td>Cabernet-S.</td>
<td>39</td>
<td>34.8</td>
<td>10.4</td>
<td>0.98***</td>
<td>1.83 ns</td>
<td>0.92***</td>
<td>32.4</td>
<td>0.96</td>
</tr>
<tr>
<td>Touriga N.</td>
<td>108</td>
<td>139.5</td>
<td>15.5</td>
<td>0.98***</td>
<td>34.3*</td>
<td>0.82***</td>
<td>126.9</td>
<td>0.90</td>
</tr>
<tr>
<td>Jaen</td>
<td>143</td>
<td>104.3</td>
<td>9.1</td>
<td>0.98***</td>
<td>24.4 ns</td>
<td>0.91***</td>
<td>114.1</td>
<td>0.96</td>
</tr>
<tr>
<td>Combined</td>
<td>446</td>
<td>65.4</td>
<td>6.1</td>
<td>0.97***</td>
<td>41.0***</td>
<td>0.89***</td>
<td>134.4</td>
<td>0.93</td>
</tr>
</tbody>
</table>

¹ t-test for intercept = 0; ² t-test for slope = 1; ns - not significant; *, **, *** P <0.05, 0.01 and 0.001 respectively.
able for the primary leaf area model. The reasons are (1) the influence of internode length on the shoot length/shoot leaf area relationship which is variety and vigour dependent (HUGLIN and SCHNEIDER 1998); (2) the change in the leaf area/shoot length ratio along the season (MABROUK and CARBONNEAU 1996); (3) the effect of trimming on the shoot length/shoot leaf area relationship; after trimming the primary shoot length remains constant but individual leaves can still grow; (4) the reduction of leaf area caused by defoliation or physical damage does not influence shoot length. These results show that models based on shoot length as the sole estimator of primary leaf area, as proposed by MABROUK and CARBONNEAU (1996), should be used with caution in situations where trimming or leaf removal are common practices or where natural defoliation is expected.

The three physical variables included in the model can be measured in a fast and easy way in the field. Yet, while the number of leaves can be obtained straightforward, the decision which is the largest and smallest leaf is a bit more laborious. It is desirable to be able to identify these leaves while counting them. While for the smallest leaf is needed only to compare the basal and apical leaves (Fig. 3 B), the choice of the largest leaf is complicated by the frequency distribution shown in Fig. 3 A. The next step is the measurement of the respective leaf veins length and/or width in order to indirectly estimate their leaf area. This estimation can be done using one of the multiple equations proposed in literature like e.g. the sum of the two lateral veins (CARBONNEAU 1976 a, LOPES and PINTO 2000) or the length of the leaf midvein and the maximum leaf width (SEPÚLVEDA and KLEIJN 1983, ELSNER and JUBB 1988). Since these empirical relations may differ between varieties it has to be set before by direct leaf area measurements.

For lateral leaf area estimation the elected model is based on two variables, one measured and one calculated but it is also the variable MLA2 which explains most of the lateral leaf area variability. These results show that the mean leaf area per shoot is the most important variable to explain shoot leaf area variability independently of the type of shoots. Inclusion of Ls in the model shows that the area of the largest lateral leaf is also an important variable to explain lateral leaf area variability as reported previously (LOPES and PINTO 2000). The low partial R² indicates however, that the model will not lose too much accuracy without L2, but since this variable is already necessary for the calculation of MLA2, no extra field work is needed.

In both models no significant interactions between the explanatory variables and the sampling date were found, enabling their applicability all along the season. Likewise the obtained models allow a separated estimation of primary and lateral leaf area which is an important feature enabling the calculation of the lateral leaf area proportion, a very important growth index (CHAMPAGNOL 1984).

Validation of the models showed a very good fit both for primary and lateral leaf area. In both models, the modeling efficiency, an overall measure of goodness of fit, was always ≥ 0.90 indicating a very good fit (MAYER and BUTLER 1993). However some differences remain in model performance. While for the area of primary leaves all data sets presented a MAE%E < 10 %, the upper limit of acceptability suggested by KLEIJNEN (1987), for lateral leaf area the Touriga Nacional data set presented a MAE%E higher than this indicating a lower goodness of fit. The regression analysis between observed and estimated values also indicate a good fit although all fitting lines have slopes significantly different from 1 indicating that the models underestimate primary leaf area and overestimate lateral leaf area by a multiplicative factor.

Acknowledgments

We thank JOSÉ VICENTE-PAULO (PhD student) for his help with Aragonez data collection. We also thank our colleagues VANDA PEDROSO and SERGIO MARTINS from the Centro de Estudos Vitivinícolas do Dão and Isabel Andrade (PhD student) for the validation datasets from the Dão wine region. The contribution of Professor St’Aubyn on statistical data analysis is gratefully acknowledged.

References


Received October 29, 2004