Image analysis-based modelling for flower number estimation in grapevine

Borja Millan*, Arturo Aquino, María P Diago and Javier Tardaguila

Instituto de Ciencias de la Vid y del Vino (University of La Rioja, CSIC, Gobierno de La Rioja),

Finca La Grajera, Ctra. De Burgos Km. 6. 26007 Logroño, La Rioja, Spain

*Corresponding author: Borja Millan, telephone +34 941 894980,
fax +34 941 899728. e-mail: borja.millanp@unirioja.es

Abstract

BACKGROUND: Grapevine flower number per inflorescence provides valuable information that can be used for assessing yield. Considerable research has been conducted at developing a technological tool, based on image analysis and predictive modelling. However, it has never been evaluated the behaviour of variety-independent predictive models and yield prediction capabilities on a wide set of varieties.

RESULTS: Inflorescence images from 11 grapevine *Vitis vinifera* L. varieties were acquired under field conditions. The flower number per inflorescence and the flower number visible in the images were calculated manually, and automatically using an image analysis algorithm. These datasets were used to calibrate and evaluate the behaviour of two linear (single-variable and multivariable) and a non-linear variety-independent model. As a result, the integrated tool composed of the image analysis algorithm and the non-linear approach showed the highest performance and robustness (*RPD*=8.32, *RMSE*=37.1). The yield estimation capabilities of the flower number in conjunction with fruit set rate (R^2=0.79) and average berry weight (R^2=0.91) were also tested.

CONCLUSION: This study proves the accuracy of flower number per inflorescence estimation using image analysis algorithm and a non-linear model that is generally applicable to different grapevine varieties. This provides a fast, non-invasive and reliable tool for estimation of yield at harvest.

Keywords: fruit set rate; yield prediction; computer vision; flowering; multi variety linear models; non-linear models;

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jsfa.7797

Introduction

The number of flowers per inflorescence is a relevant agronomical parameter, necessary to estimate the fruit set rate (number of flowers that become berries), that allows to perform early fruit yield estimation[1] and quality assessment[2]. However, practically speaking, the calculation of the flower number per inflorescence has been troublesome due to strong operational limitations, as it has been traditionally assessed manually[3]. Consequently, several methods for simplifying the flower counting procedure have been proposed[4-7]. Poni *et al.*[4] used inflorescence images to manually count the number of the visible flowers on photo prints, when recent publications[5-7] relied on image analysis algorithms for automatizing the process. Nevertheless, these solutions only provide partial information, since an inflorescence is a 3-dimensional structure in which not all the flowers will be visible in an image due to occlusions from other flowers or inflorescence structure. Therefore, in order to provide an estimation of the total flower number in inflorescence images, besides being able to count visible flowers, a predictive model is also mandatory. Such a model, taking as input the information extracted from images, has to be able to predict the total amount in the inflorescences. To this respect, investigations developed in previous studies have shown promising results. However, certain key aspects still remain inconclusive and require further investigation.

Previous developments in predictive models have been carried out on a limited number of varieties (4 in the best case[5]). It is well known that inflorescences from different varieties may have distinctive structural features, what argues for the need of additional investigation considering a wider set of varieties to confirm that the concept is generalizable. On the other hand, and with the previous limitation, several modelling approaches have been explored in the literature. Poni *et al.*[4] were the first considering the use of linear variety-dependant models. This approach was also used by Diago *et al.*[6], who additionally proposed the use of a unique, global, variety-independent linear model for all varieties due to the outstanding and obvious applicative advantages it provides. Comparing both strategies, they found poorer performance for the variety-independent global solution, although this outcome was latter contradicted by a revision carried out by Aquino *et al.*[7]. Moreover, these authors also tested the performance of a non-linear variety-independent model enriched with certain features extracted from images. The obtained results were promising, but conclusions limited because of the reduced dataset used. This was acknowledged by the authors, who stated that further research was necessary before accepting the increased complexity derived from the use of a non-linear model. In addition, all the reviewed investigations are motivated within the frame of the development of an integrated tool composed of an image analysis algorithm and a predictive model for the non-invasive and in-field estimation of the flower number per inflorescence. However, despite these two 'logical' blocks have been separately evaluated as reviewed before, their combination, as an integral solution, remains untested and, therefore, its suitability potentially unknown. Additionally, the capability to estimate the yield using the number of flowers per inflorescence in conjunction with other indicators like fruit set has not been evaluated. The absence of an easy-to-use and non-invasive tool for grapevine flower estimation has prevented from performing extensive studies on this topic.

This paper provides a multi-focal and wide study, considering 11 grapevine varieties, on the non-invasive and in-field estimation of the number of flowers per inflorescence. To this end, non-linear and linear variety-independent models were developed, tested and compared in terms of accuracy, reliability and generalization potential. Moreover, their behaviour was also assessed and discussed when they performed jointly with an evolved version for stability improvement of the image analysis algorithm presented by Diago *et al.*[6]. Yield estimation's performance from the number of flowers alone and in conjunction with the berry set rate and berry weight was appraised. This evaluation scheme allowed to test an integrated tool for flower number estimation for the first time on a wide set of grapevine varieties.

Material and methods

Image acquisition

Inflorescence images from 11 varieties of *Vitis vinifera* L. (Viognier, Verdejo, Touriga Nacional, Tempranillo, Syrah, Riesling, Pinot Noir, Grenache, Cabernet Sauvignon, Albariño and Airen) were acquired in a commercial nursery vineyard located in Falces (Navarra, Spain) on the 21st of May 2014. The images were taken at flowering. A characterization of its colour, inflorescence size and the phenological stage at the time of imaging using the BBCH scale described by Lorenz *et al.*[7] is showed in Table 1.

For each variety, 12 inflorescence images were taken placing a black cardboard behind the inflorescence for background homogenization. For image capture, a Nikon D5300 digital reflex camera (Nikon corp., Tokyo, Japan) equipped with a Sigma 50 mm F2.8 macro (Sigma corp., Kanagawa, Japan) was used. RGB images were saved at a resolution of 24 Mpx (6000x4000 pixels), 8 bits per channel and with uncontrolled illumination. No tripod was used. The distance between the camera and the inflorescence was not fixed, but within the range of 30-50 cm. The flowers from these inflorescences were manually detached from the rachis and counted. In order to obtain data for the yield estimation, another dataset consisting on 87 inflorescence images from 7 varieties (Viognier, Verdejo, Tempranillo, Syrah, Riesling, Cabernet Sauvignon and Airen) were obtained capturing the images in the same way as previously described. These inflorescences were not cut from the vine, allowing its fully development until harvest, when the clusters were picked, and berries manually destemmed, counted and weighted.

Image analysis algorithm for flower detection

The algorithm for flower detection in the inflorescence images was implemented in Matlab (Matlab R2010b, Mathworks, Natick, MA, USA) for processing batches of images in a fully automated way. This was an evolved version for stability improvement of the algorithm described in Diago *et al.*[6]

As a pre-processing step, all the images were downscaled to a 9Mpx resolution for computational workload optimization and noise reduction. Additionally, in order to decouple illumination and colour information in the images, they were converted into the CIELAB colour space (CIE 1976 $L^*a^*b^*$)[8,9].

The first step consisted on the ROI (region of interest) extraction, which was performed by binarization on the b^* channel of the CIELAB colour space for separating the inflorescence

from the background. The binarization threshold was set using the Otsu's[10] method, which enables its calculation in an automated, reliable and illumination-independent way. For that, Otsu's method assumes that the image contains 2 types of pixels corresponding to different classes (foreground and background) and that the intensity of these pixels distributes in a bi-modal histogram. Consequently, the threshold value is calculated as the one providing the optimum separation of the 2 classes.

The second step involved the detection of flower candidates, and it was based on finding the local maximum of illumination in every flower. This maximum is generated by the light reflection on the surface of the "quasi spherical" shape of the flower, which generates a maximum intensity where reflection takes place. These reflection spots were identified using a morphological transformation called h-maxima[11], as described in Diago *et al*[6], and considered as flower candidates.

The third step consisted on false positive removal by discarding those flower candidates generated by reflection spots not caused by flowers. Indeed, the higher intensity spots are usually associated to flowers, but some of them may originate from reflection points in the background or rachis. Therefore, for false positive filtering, three criteria using the following descriptors were sequentially applied:

- Size: the reflection produced by flowers was expected to have a similar size. This descriptor quantifies the size of the flower candidates in terms of number of pixels.
- Distance: this descriptor was formulated by taking the assumption that flowers form clusters in the inflorescence, and that the distance between flowers within a same cluster is similar. Also assuming that a cluster is composed by, at least, five flowers, the distances between any flower within a cluster and its five closest neighbours must be short and similar. Hence, the distance descriptor was calculated for every flower candidate as the fifth closest distance between the one under evaluation and the remaining candidates.
- Shape: the bright spots produced by flowers described a circular pattern. This descriptor evaluates the circularity of the flower candidates by means of the quotient between the minor and major axis.

Once the descriptors were calculated for a given image, the filtering procedure relied on processing their histograms of values. Concretely, a threshold for every histogram was set and all the flower candidates with descriptor values that set over the threshold were discarded. For deciding a threshold value, the histograms were assumed to describe normal distributions. With this assumption, the thresholds were calculated as twice the standard deviation of the distributions. In the original work, Diago *et al*.[6] found this point by implementing a process consisting on finding the first local minimum after the global maximum of the smoothed histogram by means of Savitzky-Golay filtering[12]. The new approach implemented in this study improved the stability in threshold calculation, as the first local minimum placed after the global maximum in the histogram not always corresponded to the end of the flower data cluster, but sometimes to discontinuities in the histogram related to a discrete dataset.

Model generation

Three different variety-independent models involving linear and non-linear approaches were considered. The studied models are described as follows:

- Single variable linear model: a unique linear model, valid for all varieties, was built using as input information the number of flowers visible in an image. For its generation, 55 images from all varieties were selected, processed and results used to obtain the best-fitted linear function. The remaining 77 images were saved for external validation of the model.

- Multivariable linear model: the linear model previously defined only used the number of flowers visible in the image as input for predictions. However, there might be other inherent and distinctive structural features in the inflorescence influencing the total flower number. From their study and characterisation, a set of five mathematical descriptors was generated:

    1. Number of flowers visible in the image.
    2. ROI area: It was calculated as the number of pixels in the image belonging to the ROI.
    3. Flower radius estimation: It was determined as the average of the minimum distances among flowers (this was assumed as equivalent to the flower diameter estimation) divided by two.
    4. Flower density: It was defined as the number of flowers divided by the ROI area.
    5. Flower area: It was calculated as the area occupied by a flower in logarithm to base 10 scale. The flower area was estimated by assuming the flower to be circular in shape, and using the previous radius estimation.

    The model was generated using 55 images and evaluated with the remaining 77 images.

- Non-linear model: The five descriptors previously described were used to characterize a non-linear model obtained with a multilayer feed-forward backpropagation neural network. Therefore, the neural network had five input neurons fed by the above defined descriptors, one hidden layer with two neurons and one output; the transfer function was set to linear. The neural network was trained using data from the same set of 55 images used for obtaining the linear model, and externally validated as described below using data from the remaining 77 images.

Performance analysis

A three-step evaluation scheme was designed to study: (a) the accuracy of the image analysis algorithm, (b) the behaviour of the developed models and (c) their combined performance.

Details concerning the aims and structure of each of these three sections are described below.

Evaluation of the image analysis algorithm for flower detection

The new version for improved stability of the image analysis algorithm was evaluated on the set of 132 images whose inflorescences were cut and manually destemmed. For that, the images were analysed with the algorithm and the results compared to reference values. These

reference values were obtained by manually labelling flowers on the 132 images making use of a custom application developed in Matlab 2010 (Figure 1). The set composed of these manually segmented images used as reference is referred to as gold standard set hereafter. Then, the algorithm outputs and the gold standard references were compared, with the help of the same *ad-hoc* Matlab-based software, for calculating true positives, false positives and false negatives following the definitions below:

- True positive (*TP*): It is an automatically detected flower corresponding to an actual flower in the gold standard set.
- False positive (*FP*): It is an automatically detected flower not corresponding to an actual flower in the gold standard set. A redundant *TP* was also considered as *FP*.
- False negative (*FN*): It is an actual flower, labelled in the gold standard set, which was not found by the image analysis algorithm.

Finally, the Recall (*RC*) and Precision (*PR*) metrics were used for evaluating the quality of each analysed image as follows:

$$RC = \frac{TP}{TP + FN} \quad (1)$$

where Recall provides the percentage of actual flowers detected;

$$PR = \frac{TP}{TP + FP} \quad (2)$$

where Precision indicates the percentage of flowers correctly assessed.

Evaluation of the predictive potential of the estimation models

The goal of this evaluation was to strictly assess the theoretical, hence maximum, estimation potential of the developed linear and non-linear models. To be able to confidently make solid conclusions in this sense, the accuracy of the data supplied as input to the estimation models had to be guaranteed. Under this precept, the outputs produced by the image analysis algorithm were not suitable for this evaluation, as they may introduce external errors that can distort and impoverish the results of the study. As an alternative, the gold standard set previously described and obtained by manually labelling flowers on the 132 images was used as input for model building (as described in the Model Generation section) and for external validation. Indeed, following this procedure, higher certainty was achieved on the data correctness.

On the other hand, estimation results produced by the models had to be compared to reference data. With this goal, all the photographed inflorescences were individually collected, destemmed, flowers manually counted and results were annotated. Thus, the actual flower number values per inflorescence were compared to the outcomes yielded by the models.

The quality of the estimations produced by the models was assessed using the determination coefficient (R^2), the root-mean-square error (*RMSE*) and the ratio of performance to deviation (*RPD*). Mathematically *RMSE* is described as follows

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\widehat{Fl}_i^v - Fl_i^v)^2}{n}} \quad (3)$$

where $\widehat{Fl}_i^v$ and Fl_i^v are the predicted and actual flower number values for the *ith* image in the validation set, respectively. Additionally, *RPD* is defined as

$$RPD = \frac{\sigma^t}{RMSE} \quad (4)$$

being σ^t the standard deviation of the actual flower number values in the training set.

RMSE offers an absolute value of the prediction error for a certain model. However, since *RMSE* does not take into account the sparsity and the range of values of the distribution, it might not be completely descriptive on its own. Therefore, *RPD* was also introduced[13] to complement the information provided by the *RMSE*, as it calibrates the obtained *RMSE* to the standard deviation of the values in the training set. According to the literature, a model with a *RPD* value between 2 and 2.5 indicates that coarse quantitative predictions are possible, a value between 2.5 and 3 corresponds to good prediction accuracy; values equal or greater than 3 indicate excellent prediction accuracy[14].

Evaluation of the performance of an integrated tool for grapevine flower estimation

The performance of an integrated tool, composed of the image analysis algorithm and the estimation models described here, for the fully automated, non-invasive and in-field estimation of the flower number per grapevine inflorescence was assessed. For analysing the performance of the linear and non-linear models when interacting with the image analysis algorithm, a version of the integrated tool was created for each model. Then, the two versions including the two different models were evaluated and compared as described in the previous section; this is, in terms of R^2, $RMSE$ and RPD.

Evaluation of the predictive capability of flower number data for yield estimation

The possibility to predict the yield from the flower number has a great application and impact because of its early stage estimation. To test its potential accuracy, the number of flowers estimated using the non-linear model and the image analysis algorithm was used to estimate the number of flowers of the second set of 87 inflorescence images. The inflorescences were photographed without cutting them from the vine and allowing their fully development until harvest, when they were collected.

The number of flowers and berries are related by the *Fruit set rate*. This rate is known to vary from one variety to another and also depends on the weather conditions at berry set and the carbohydrate availability of each plant. To allow a proper berry number per cluster estimation, *Fruit set rate* was calculated separately for every variety under study (Viognier, Verdejo, Tempranillo, Syrah, Riesling, Cabernet Sauvignon and Airen). It was determined as the total number of berries in the clusters of that variety (obtained by manual destemming the berries

from the cluster) divided by the total number of flowers in the inflorescences (estimated with the image analysis algorithm and the non-linear model):

$$Fruit\ set\ rate^v = \frac{\sum_{i=1}^{n} Bn_i^v}{\sum_{i=1}^{n} \widehat{Fl}_i^v} \quad (5)$$

where Bn_i^v is the actual berry number for the *ith* cluster in the set.

This calculation allowed to obtain a fruit set rate for every variety under study. This value was used to obtain the *Estimated cluster berry number* as follows:

$$Estimated\ cluster\ berry\ number_i^v = \widehat{Fl}_i^v * Fruit\ set\ rate^v \quad (6)$$

Similarly for yield estimation, *Average berry weight* was calculated as the weight of the berries in the clusters divided by the number of berries in the clusters of that variety obtained by manual destemming:

$$Average\ berry\ weight^v = \frac{\sum_{i=1}^{n} Bw_i^v}{\sum_{i=1}^{n} Bn_i^v} \quad (7)$$

where Bw_i^v is the total weight of the berries in the *ith* cluster in the set.

From this it is possible to obtain the *Estimated yield* as follows:

$$Estimated\ yield_i^v = \widehat{Fl}_i^v * Fruit\ set\ rate^v * Average\ berry\ weight^v \quad (8)$$

Results and discussion

Performance evaluation of the image analysis algorithm for flower detection

Table 2 shows the obtained average and standard deviation values of Recall ($\overline{RC}$ and σ_{RC}, respectively) and Precision ($\overline{PR}$ and σ_{PR}, respectively), detailed per variety and considering all images as a whole. As it can be observed, the produced results can be considered satisfactory and consistent among varieties attending to the measured averages and standard deviations. The $\overline{RC}$ values were all over 0.80 (except for Airen with $\overline{RC}$=0.79) and in conjunction with $\overline{PR}$ values higher that 0.75 validates the general applicability of the image analysis algorithm in flower detection. With this premise, more accurate results for certain varieties can also be observed when carefully comparing partial outcomes (for example, checking differences between Airen and Grenache in Table 2). This effect can be attributed to dissimilarities in structure and phenological stage of the inflorescences among varieties at the time of image acquisition. Indeed, these differences were present since all the images were taken the same day, and the varieties for this study were carefully selected to significantly differ in terms of

dimensions of the inflorescence, size of the flowers, colour or phenological development as it can be observed in Figure 2.

The measured results even outperformed for recall and provided similar precision results to those reported by Diago *et al.*[6] with the original algorithm's version, tested on three varieties and 15 images. These differences cannot be attributed to the modifications applied to the new implementation, since they were aimed at improving stability in the calculation of the required thresholds for the filtering criteria as previously described. However, the results of this wide and exhaustive study outstandingly increased the confidence in the general and successful applicability of the methodology.

Figure 3-(A) shows an original RGB image of an inflorescence of cv. Airen taken in the vineyard, and image (B) represents the flower detection result, with the identified flowers being signalled with red dots. This inflorescence image had 530 flowers (manually assessed), and the image analysis algorithm correctly identified 412 of them, producing 31 *FP*s. With these figures, the obtained Precision and Recall values were 0.93 and 0.78, respectively.

Evaluation of the predictive potential of the estimation models

As it was previously discussed, linear and non-linear models were generated and tested in this study using gold standard data created by manually segmenting flowers on the images. Concretely, from the gold standard set of 132 manually-labelled images, 55 of them were kept for model training. Data from these images, along with their corresponding actual flower number manually and destructively counted, were used as input for model calibration. The remaining 77 samples were used for model external validation.

The relationship between the actual flower number per inflorescence and the predictions yielded by the linear and non-linear models can be graphically analysed in Figure 4. Moreover, Table 3 includes the obtained results in terms of the defined metrics and detailed per grapevine variety. Attending to all metrics, it can be confirmed great performance for all of the approaches, with overall results virtually equivalent. In fact, attending to the *RPD* value given by the models, their performance can be qualified as excellent as stated in the literature (*RPD* > 3); this qualification was consistent for all varieties. Focusing on partial results per variety, equivalent results and slightly better for one approach can be found, but there was not any recognizable pattern justifying this behaviour.

The coefficient of determination (R^2) values measured for both the linear and non-linear models developed in this work were comparable to those obtained by Aquino *et al.*[7]. This fact confirms the appropriate general behaviour of the models for flower estimation outlined in that work. However, when attending to the *RMSE*, which accurately measures the magnitude of the predictive mistake, the differences outstandingly increased, being figures measured by Aquino *et al.*[7] considerably worse than those obtained in this study (38.0 vs 84 of RMSE for the lineal model; 38.4 vs 58 for the non-linear model), meaning that the deviation from the prediction in Aquino *et al.*[7] doubles the obtained in this study, that is, in any case an excellent prediction due to its RPD values of 8.12 and 8.04 respectively. This discovery, which indicates higher predictive potential than that found by Aquino *et al.*[7], may be explained by the reduced

datasets used by those authors for training and testing (they were composed of 28 and 20 samples, respectively). In fact, this issue was discussed by the authors as: 'in the authors' opinion, this line needs further research before arriving at definitive conclusions. In effect, suitability of the developed feature space needs to be verified on a wider range of varieties'.

On the other hand, the analogous performance showed by the three models in this study suggests no benefits in using more descriptors in the linear model or the non-linear approach. This outcome reveals that the existent relationship between the flower number visible on images and the actual flower number per inflorescence is predominantly linear and with little influence of other variables associated to variability between cultivars, thus not existing a strong varietal dependence. This shows some discrepancy with the findings of Diago et al.[6], who stated that a variety-independent estimation linear model would not be accurate enough (experiments were carried out on three grapevines varieties). This outcome is remarkable from a practical point of view, since it avoids the requirement of using individually tuned models per variety. Indeed, variety-dependant models lead to a lack of usability for in-field appplications[15] when compared to global models.

Performance evaluation of an integrated autonomous tool for flower estimation

The evaluation of an integrated tool, composed of the image analysis algorithm and one of the developed models, aimed at estimating the flower number per grapevine inflorescence in the field in an automated and non-invasive way was attempted.

Figure 5 illustrates the relationship found between the predicted and actual flower number per inflorescence when using the integrated tool. Additionally, Table 4 specifies the measured results in terms of R^2, *RMSE* and *RPD*, detailed per grapevine variety. Attending to the results of the single variable linear model, in spite of being satisfactory, a remarkable performance degradation can be identified with respect to their theoretical potential measured in the previous section. This outcome was expected, as this model only used the number of flowers visible on the images as predictive information, and the image analysis algorithm error making in detection is higher than that of a human observer. This can be corroborated by the results of the multivariable linear model. The predictions of this model are not only based on the number of flowers visible on the image, but are also influenced by four additional descriptors, which make the model more robust against mistakes introduced by the image analysis algorithm. The inclusion of these descriptors improved the results of the single variable linear model, leading to a reduction in the RMSE values from 59.9 to 46.7 as shown in Table 2. However, the performance provided by the non-linear model virtually matched to its theoretical maximum. This can be driven by mistakes introduced by the image analysis algorithm, which are not linearly related to the descriptors. Indeed, when an increase of performance was obtained by the multivariable linear model (in comparison with the single variable linear model), the non-linear one was the only one capable to correct the errors of the image analysis algorithm.

The results obtained in this study clarify uncertainties contemplated by Aquino *et al.*[7]. Concretely, these authors considered that, in spite of being promising, the suitability of the

non-linear approach should be guaranteed with a wider study including more varieties. Additionally, they stated that the gained accuracy had to be assessed to decide whether it justified the utilisation of the non-linear approach, for being a more complex implementation than the linear one. A mixed solution like the multivariable linear model that used the descriptors of the non-linear model was not tested previously to this publication. The performed analysis indicates that, in absence of external errors, the linear and non-linear approaches provide a similar performance. However, in practice, when interacting with the image analysis algorithm, the non-linear solution offered a more robust and accurate behaviour than the linear ones, even when the same descriptors were used as inputs. In any case, the simplicity and ease of computation of the multivariable linear model can offer an advantage in some applications where the computation time is a priority. Therefore, the prediction capabilities shown in the present work on a wide range of varieties by the integrated solution composed of the image analysis algorithm and the non-linear model, indicate that, in practice, this is the optimum combination for most of the use cases.

Evaluation of the predictive capability of flower number for yield estimation

As a preliminary analysis, the relationship between the number of flowers and the final crop yield was studied. To this end, the image analysis algorithm and the non-linear model were used to provide an estimation of the flower number per inflorescence. Figure 6-(A) shows that a poor relationship existed between these two variables (R^2=0.49). This can be expected because of the differences among varieties in the fruit set rate, which is the number of flowers that became berries.

The berry set rate was used to overcome the differences in the fertility between grapevine varieties. This is shown in Figure 6-(B), with a good coefficient of determination (R^2=0.79). This relationship is crucial, because it points out that a yield prediction can be done as sooner as at the flowering stage, months before harvest. To that end, it is compulsory to have an *a priori* knowledge of the berry set rate. This factor can be obtained from historical records and may be modified by external factors related to meteorological or viticultural practices. Further research is needed to stablish a protocol to obtain the berry set rate and its precision for different varieties and sites. The use of the tool presented in this publication in conjunction with others that will allow to obtain the grape berry number in a similar way that VitisFlower app does will allow to conduct the needed investigations.

Increased precision in yield prediction can be attained if the average berry weight is known, as shown in Figure 6-(C), obtaining a strong correlation (R^2=0.91), that allows an accurate prediction of the final crop yield from the number of flowers and two additional factors (the berry set rate and average berry weight), that could be obtained from historical records and adjusted over the season to account for meteorological variations.

Conclusions

This paper presents a deep study on the estimation of the flower number per inflorescence in several grapevine varieties by means of a non-invasive, automated and in-field applicable solution. The measured theoretical potential of variety-independent models indicated that,

with a unique model, the actual flower number per inflorescence can be accurately estimated from features extracted from images. This definitely discards the need of individual models per variety, what supposes outstanding applicative advantages. On the other hand, when testing an integrated tool comprising the image analysis algorithm and the three studied models, results indicated that the performance of the non-linear approach remained close to its theoretical maximum. For the case of the linear approaches, in spite of providing good results, remarkable performance declining was detected. These findings, supported by a wide study comprising eleven varieties, definitely indicated that grapevine flower estimation in the defined circumstances can be accurately achieved, additionally arguing for the exploitation of the non-linear solution. The potential capability for final yield prediction from flower number estimation was also evaluated. The number of flowers per inflorescence alone was found to provide insufficient information for yield estimation, but procures promising results when combined with fruit set rate and especially when the average berry weight is known. The computation of the berry set rate requires the assessment of the number of flowers in a not destructive way, so the use of the presented algorithms offers an important advance to this end. The integrated solution can be implemented in a smart device application, such as the existing vitisFlower (that is available for Android based devices through android applications market), to provide the grape industry with a fast, non-invasive, reliable and inexpensive tool to automatically estimate the number of flowers per inflorescence under field conditions.

References

1 May P, *Flowering and fruitset in grapevines.* Phylloxera and Grape Industry Board of South Australia, Adelaide, Australia, (2004).

2 Matthews M and Nuzzo V, Berry size and yield paradigms on grapes and wines quality. *Acta Hort* **754**:423-436 (2007).

3 Guilpart N, Metay A and Gary C, Grapevine bud fertility and number of berries per bunch are determined by water and nitrogen stress around flowering in the previous year. *Eur J Agron* **54**:9-20 (2014).

4 Poni S, Casalini L, Bernizzoni F, Civardi S and Intrieri C, Effects of Early Defoliation on Shoot Photosynthesis, Yield Components, and Grape Composition. *American Journal of Enology and Viticulture* **57**:397-407 (2006).

5 Aquino A, Millan B, Gutiérrez S and Tardáguila J, Grapevine flower estimation by applying artificial vision techniques on images with uncontrolled scene and multi-model analysis. *Comput Electron Agric* **119**:92-104 (2015).

6 Diago MP, Sanz-Garcia A, Millan B, Blasco J and Tardaguila J, Assessment of flower number per inflorescence in grapevine by image analysis under field conditions. *J Sci Food Agric* (2014).

7 Lorenz DH, Eichhorn KW, Bleiholder H, Klose R, Meier U and Weber E, Growth Stages of the Grapevine: Phenological growth stages of the grapevine (Vitis vinifera L. ssp. vinifera) - Codes

and descriptions according to the extended BBCH scale. *Australian Journal of Grape and Wine Research* **1**:100-103 (1995).

8 Schanda J, *Colorimetry : understanding the CIE system.* CIE/Commission Internationale de L'eclairage, Vienna (Austria), (2007).

9 Connolly C and Fleiss T, A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *Image Processing, IEEE Transactions on* **6**:1046-1048 (1997).

10 Nobuyuki O, A Threshold Selection Method from Gray-Level Histograms. *Systems, Man and Cybernetics, IEEE Transactions on* **9**:62-66 (1979).

11 Soille P, *Morphological image analysis: principles and applications.* Springer-Verlag New York, Inc., (2003).

12 Savitzky A and Golay MJE, Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal Chem* **36**:1627-1639 (1964).

13 Williams P and Norris K, *Near-infrared technology in the agricultural and food industries.* American Association of Cereal Chemists, Inc., St. Paul, Minnesota, USA, (1987).

14 Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI and Lammertyn J, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol Technol* **46**:99-118 (2007).

15 Aquino A, Millan B, Gaston D, Diago M and Tardaguila J, vitisFlower®: Development and Testing of a Novel Android-Smartphone Application for Assessing the Number of Grapevine Flowers per Inflorescence Using Artificial Vision Techniques. *Sensors* **15**:21204 (2015).

Tables

Table 1: Characterization of the grapevine variety colour, inflorescence size and phenological stage at the time of image acquisition.

Grapevine variety	Colour	Inflorescence size (average #flowers/inflorescence)	Phenological stage
Airén	White	738	53-55
Albariño	White	277	55-57
Cabernet Sauvignon	Red	419	53-55
Grenache	Red	613	55-57
Pinot Noir	Red	181	55-57
Riesling	White	237	55-57
Syrah	Red	169	53-55
Tempranillo	Red	242	53-55
Touriga Nacional	Red	192	55
Verdejo	White	483	53-55
Viognier	White	228	55

[a]Phenological stage at which inflorescence images were taken. The numbers refer to the phenological growth stages according to the BBCH-Identification Key of Grapevine in BBCH (Lorenz et al., 1994). BBCH 53: Inflorescence clearly visible; 55: Inflorescence swelling: flowers pressed together; 57: Flowers separate; inflorescence developed.

Table 2: Results obtained with image analysis algorithm in terms of average and standard deviation values of Recall ($\overline{RC}$, σ_{RC}) and Precision ($\overline{PR}$, σ_{PR}). Figures are detailed per variety (n=12).

Variety	$\overline{RC}$	σ_{RC}	$\overline{PR}$	σ_{PR}
Viognier	0.86	0.03	0.82	0.05
Verdejo	0.89	0.03	0.81	0.05
Touriga Nacional	0.89	0.04	0.88	0.06
Tempranillo	0.89	0.03	0.87	0.03
Syrah	0.81	0.05	0.81	0.07
Riesling	0.83	0.04	0.87	0.05
Pinot Noir	0.86	0.04	0.79	0.09
Grenache	0.87	0.04	0.86	0.05
Cabernet Sauvignon	0.86	0.05	0.89	0.05
Albariño	0.89	0.05	0.76	0.08
Airen	0.79	0.06	0.90	0.03
Overall	0.86	0.05	0.84	0.07

Table 3: Results yielded by the linear (single variable and multivariable) and non-linear models to estimate the actual flower number per inflorescence from features manually extracted from images (n=77)

Variety	Single variable linear model			Multivariable linear model			Non-linear model		
	R^2	RMSE[a]	RPD[b]	R^2	RMSE[a]	RPD[b]	R^2	RMSE[a]	RPD[b]
Viognier	0.88	39.2	7.86	0.89	39.6	7.79	0.91	22.8	13.56
Verdejo	0.97	38.6	7.98	0.96	40.4	7.64	0.97	52.4	5.89
Touriga Nacional	0.70	58.0	5.32	0.72	62.5	4.94	0.75	51.3	6.01
Tempranillo	0.99	17.3	17.80	0.99	16.6	18.55	0.99	19.5	15.81
Syrah	0.97	28.9	10.66	0.97	30.3	10.18	0.95	35.0	8.82
Riesling	0.94	25.9	11.92	0.94	27.2	11.35	0.95	17.9	17.29
Pinot Noir	0.82	26.7	11.57	0.84	24.7	12.50	0.86	41.6	7.42
Garnacha	0.88	58.9	5.24	0.95	57.9	5.33	0.91	40.1	7.68
Cabernet Sauvignon	0.99	26.3	11.74	0.99	24.3	12.72	0.99	29.3	10.53
Albariño	0.68	26.4	11.67	0.78	22.6	13.67	0.83	33.3	9.26
Airen	0.98	43.4	7.12	0.98	45.1	6.84	1.00	43.3	7.13
Global	0.96	38.0	8.12	0.96	38.5	8.01	0.96	38.4	8.04

[a]: root-mean-square error; represents the average root deviation between predicted and observed values.
[b]: ratio of performance-to-deviation; is the ratio between the standard deviation of the sample set used to train a model and the RMSE measured for that model.

Table 4: Results yielded by the linear (single variable and multivariable) and non-linear models in the estimation of the total flower number per inflorescence from features extracted using image analysis algorithm (n=77).

Variety	Single variable linear model			Multivariable linear model			Non-linear model		
	R^2	RMSE[a]	RPD[b]	R^2	RMSE[a]	RPD[b]	R^2	RMSE[a]	RPD[b]
Viognier	0.70	62.4	4.94	0.75	23.6	13.08	0.89	15.6	19.74
Verdejo	0.95	83.3	3.70	0.96	52.4	5.89	0.95	52.4	5.89
Touriga Nacional	0.34	67.2	4.59	0.19	32.5	9.50	0.52	37.4	8.25
Tempranillo	0.94	20.4	15.09	0.95	10.2	30.22	0.99	10.2	30.22
Syrah	0.76	36.5	8.46	0.84	28.8	10.70	0.86	32.2	9.59
Riesling	0.82	29.7	10.38	0.72	39.1	7.90	0.81	32.6	9.46
Pinot Noir	0.45	48.2	6.39	0.17	38.5	8.01	0.43	44.8	6.88
Garnacha	0.79	74.5	4.14	0.84	47.8	6.46	0.91	55.6	5.55
Cabernet Sauvignon	0.98	24.1	12.79	0.99	21.4	14.45	0.99	21.4	14.45
Albariño	0.58	77.3	3.99	0.65	30.8	10.03	0.79	22.4	13.80
Airen	0.37	98.2	3.14	0.84	126.9	2.43	0.85	60.6	5.09
Global	0.91	59.9	5.15	0.94	46.7	6.61	0.96	37.1	8.32

[a]: root-mean-square error; represents the average root deviation between predicted and observed values.
[b]: ratio of performance-to-deviation; is the ratio between the standard deviation of the sample set used to train a model and the RMSE measured for that model.

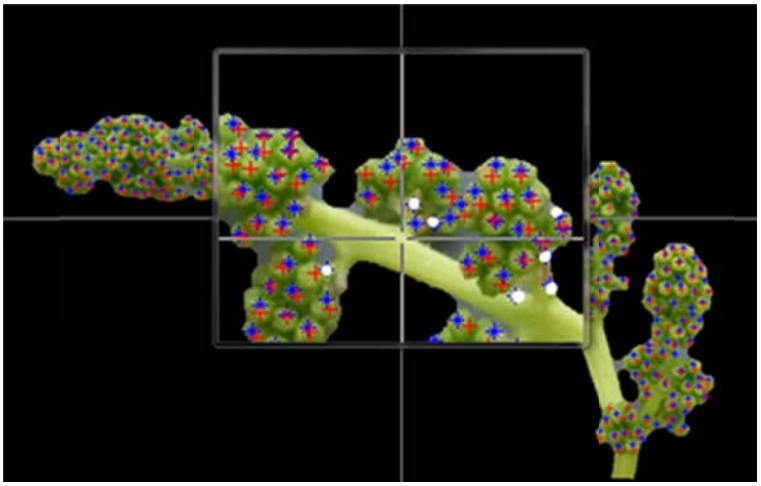

Figure 1: Screen capture of the application developed for manual ground truth generation and for evaluating the image analysis algorithm. Red crosses represent manually detected flowers, blue marks denote the automated detected flowers and white dots identify false positives produced by the algorithm.

Figure 2: Images of (a) a Tempanillo inflorescence at phenological stage BBCH 53-55, and (b) a Pinot Noir inflorescence at phenological stage BBCH 55-57. Phenological growth stages are indicated according to the BBCH-Identification Key of Grapevine in BBCH (Lorenz et al., 1994). (BBCH 53: Inflorescence clearly visible; 55: Inflorescence swelling: flowers pressed together; 57: Flowers separate; inflorescence developed).

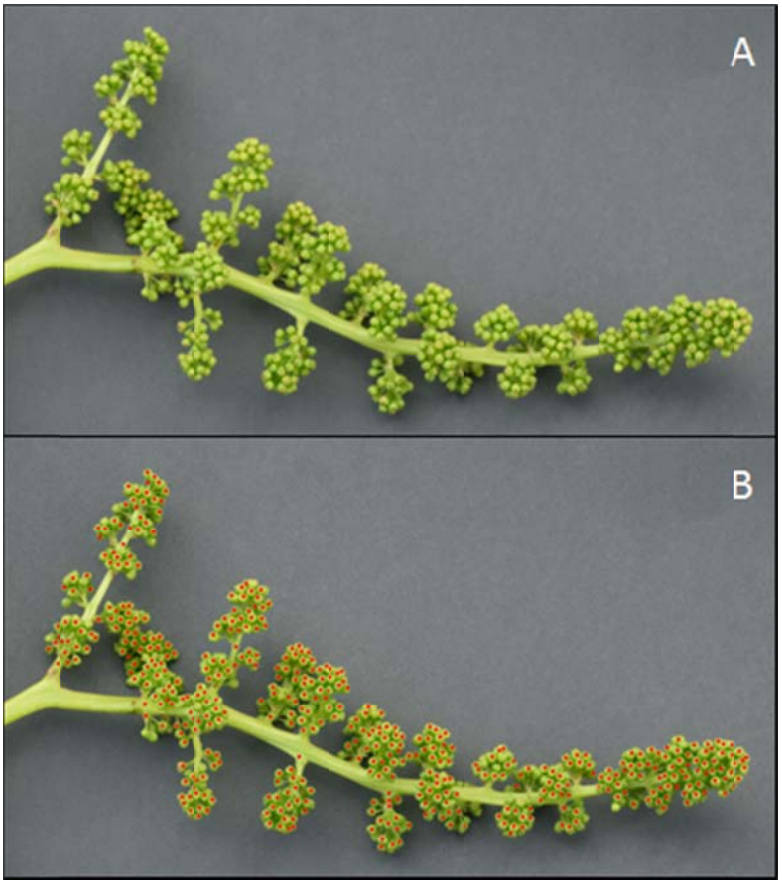

Figure 3: Illustration of the results produced by the image analysis algorithm for detecting grapevine flowers: (A) Original inflorescence image of cv. Airen taken under field conditions; (B) Flowers detected by the algorithm marked with red dots over the original image.

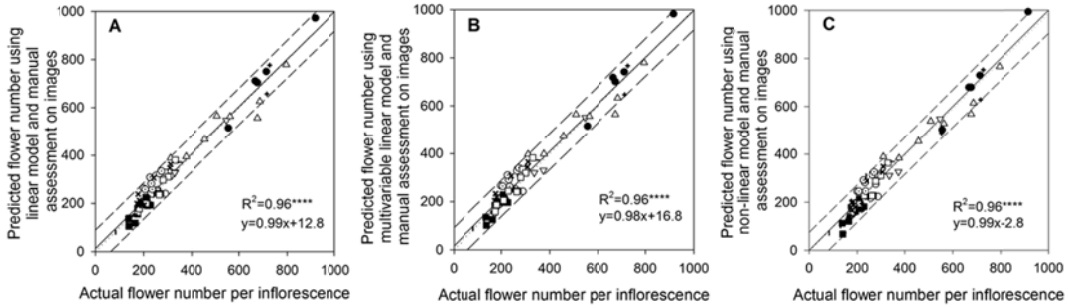

Figure 4: Graphical representation of the predicted flower number using manual counting (an operator manually tagged all the flowers in the image using a custom built computer application) vs actual flower number per inflorescence in the images of the validation set, using (A) a linear model, (B) 5 inputs multivariable linear model and (C) a nonlinear model built for 11 grapevine varieties (● Airen, ○ Albariño, ▽ Cabernet Sauvignon, △ Grenache, ■ Pinot Noir, □ Riesling, ι Syrah, – Tempranillo, ⊙ Touriga Nacional, + Verdejo, × Viognier). The

regressions were significant at p<0.001. Dotted line represents the 1:1 line and dashed lines correspond to the 95% prediction intervals. (n=77) 333x105mm

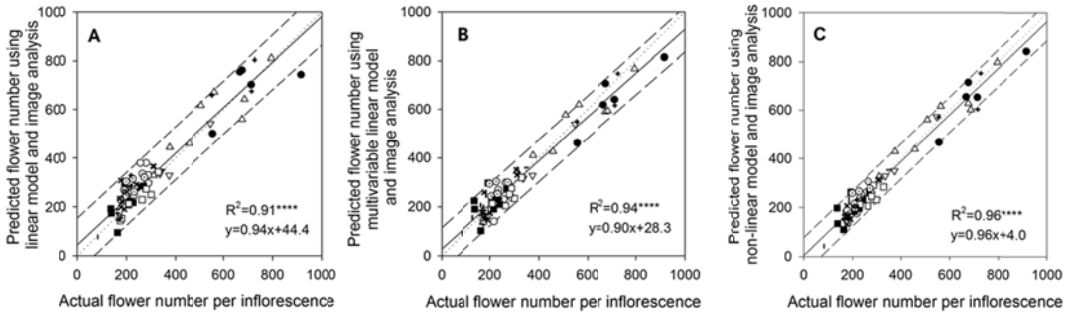

Figure 5: Graphical representation of the predicted flower number per inflorescence vs. actual flower number per inflorescence using image analysis (a fully automatized algorithm processes the images to segment the flowers) and (A) a linear model, (B) 5 inputs multivariable linear model and (C) a nonlinear model built for 11 grapevine varieties (● Airen, ○ Albariño, ▽ Cabernet Sauvignon, △ Grenache, ■ Pinot Noir, □ Riesling, ι Syrah, – Tempranillo, ⊙ Touriga Nacional, + Verdejo, × Viognier). The regressions were significant at p<0.001. Dotted line represents the 1:1 line and dashed lines correspond to the 95% prediction intervals. (n=77) 328x105mm

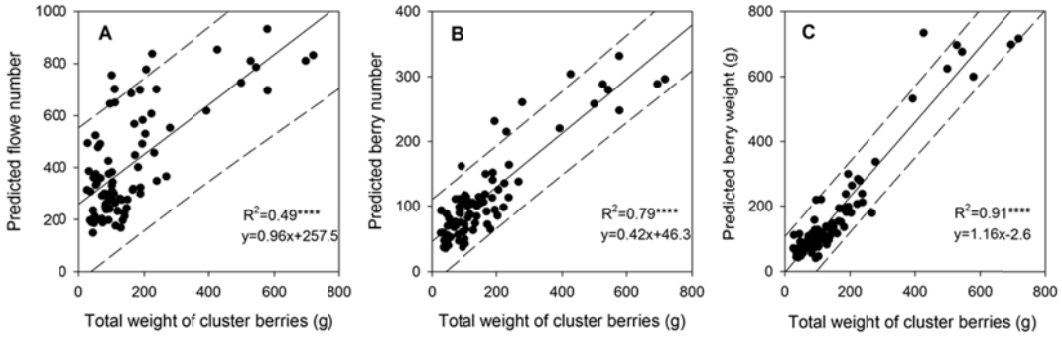

Figure 6: Graphical representation of the total weight of cluster berries vs. (A) predicted flower number per inflorescence using image analysis (a fully automatized algorithm processes the images to segment the flowers) and neural network model, (B) predicted berry number per cluster obtained using the estimation of flower number multiplied by berry set rate and (C) predicted berry weight per cluster obtained using the estimation of berry number multiplied by average berry weight. The regressions were significant at p<0.001. Dashed lines correspond to the 95% prediction intervals. (n=87) 307x105mm