REV: OPTICAL SYSTEM FOR GRAPE CROP ESTIMATION

ERIC SERRANO[1], SYLVIE ROUSSEL[2], LAURE GONTIER[1], THIERRY DUFOURCQ[1], CHRISTOPHE GUIZARD[3]

[1] ITV France Midi-Pyrénées, V'Innopôle, BP 22, 81310 Lisle/Tarn, France
eric.serrano@itvfrance.com
[2] ONDALYS, 361 rue JF Breton, BP5095, 34196 Montpellier, cedex 1, France roussel@agrometrix.com
[3]CEMAGREF, 361, rue J.F. Breton, B.P. 5095,34196 Montpellier Cedex 5, France

Extract from the proceedings of the technical seminar on innovation in Viticulture/Enology organized by the Station Régionale ITV Midi-Pyrénées on December 15th 2005 in Toulouse.

INTRODUCTION:

Wine production has always been characterized by inconsistent harvest volumes from one year to the next. One of the challenges is to provide the professionals involved in crop estimation with a reliable and fast method allowing early predictions.

Different methods have tried to predict harvest volumes for more or less extended areas. The first estimation attempts are reported in the bibliography in 1955 (Wurgler, Leyvraz et Bolay).

At the plot level, the most comprehensive studies focussed mainly on counting inflorescences, clusters per vine or berries per cluster (Murisier, 1986).

Studies carried out in Alsace (1957-1971, Huglin and Schneider) indicated that the average cluster weight was highly variable and a source of errors. Dufourcq and Serrano (2003) demonstrated that the annual variation coefficient of the average cluster weight on a same plot of Negrette reached 15%. This observation confirmed the results of Huglin et *al.* (1975) with Syrah, or of Gerbier and Remois (1977) in the Champagne area.

Also, while some of the proposed methods allowed to obtain satisfactory estimations, they have not been widely applied as they either require long-standing historical data (Schneider, 1997), or an overly demanding implementation (counting berries per cluster and per shoot, and the weight of the berries) in order to provide quality information (Booysen et Al, 1978).

This study requires to abandon measurement of the variables “number of berries per clusters” and “average berry weight” because of its cumbersome implementation.

From 1999 to 2004, the study focussed on demonstrating a correlation between the cluster volume during its development and its weight at harvest. During the first phase of the berry growth an optimal measurement period (i.e. cluster closure) has been determined to develop a mathematical correlation model (Serrano et Al, 2005).

Considering the results, a portable optical sensor was developed and tested in order to validate the method. From 2004 to 2005, the research program concentrated on validating the performance of the system regarding repeatability and reproducibility, and on generating statistical correlation models based on the volumes calculated from the images.

MATERIAL AND METHODS:

More than twenty thousand measurements of cluster volumes at different phenological stages of the vine and cluster weights at harvest were collected. This experimental database includes nineteen *Vitis vinifera* grape varieties over a period of six vintages (1999 to 2004): Cabernet Franc, Cabernet Sauvignon, Chardonnay, Colombard, Auxerois, Duras, Fer Servadou, Gamay, Grenache, Gros Manseng, Loin de l'œil, Mauzac, Merlot, Muscadelle, Negrette, Sauvignon B., Syrah, Tannat and Ugni B. For each variety, two to three plots were identified according to their level of production or the type of soil.

Initially, the shape of the cluster was compared to a cone. Fruit measurements focussed on the height and circumference of the cluster using a prototype tool allowing direct measurement.

In 2005, the volume of a cluster is simply assessed with a photography taken with a portable sensor: the REV. The image is then processed to approach the real volume of a cluster. The method consists in calculating the cluster volume based on the two dimensional image by cutting the image into horizontal slices (approximately 50). For each slice, the rectangle on the segmented surface of the cluster is calculated, and the volume of the slice is calculated using the cylindrical formula (thus assuming axial symmetry of the cluster).

A sample of 50 to 200 clusters per plot was measured during cluster setting, at the beginning of bunch closure and veraison. The clusters were ringed and identified at the first measurement and individually weighed at harvest time.

Repeatability and reproducibility tests:

A series of trials was carried out in order to assess the repeatability and reproducibility of cluster weight estimations based on the images obtained with the portable REV sensor. Twenty clusters were measured 5 times by several users.

Statistical treatments concentrated on average and intra-operator repeatability tests.

For each user, the standard deviation, the median and quartiles of the 5 measurements of the same cluster were calculated.

The reproducibility integrates the repetitions of the measurements performed by all the users.

Linear correlation model:

The models proposed are based on linear regressions, according to the equation $\hat{w} = a_o + a_1 * v$, $\hat{w}$ being the weight estimation.

Initially, the models were calculated for each grape variety, vintage and phenological stage. Then, they were validated according to the Cross-Validation statistical method (leave-one-out cross-validation). This validation method consists in generating N models by permutation, each model being calibrated on N– 1 individuals, and tested against the individual which was excluded from the model.

The relevance of the models was evaluated by the Standard Error of Cross-Validation (SECV). This error can be expressed in % with regards to the average cluster weight of the plot studied, and provides a Coefficient of Variation.

Subsequently, the models were validated by a series of independent tests. The models were calculated using the data of previous vintages and assessed according to their ability to predict the cluster weight of the 2004 vintage.

The software used to process the data was MATLAB (The Mathworks, NATIC, USA).

RESULTS AND DISCUSSION:

Development of REV – Repeatability and reproducibility tests

Figure 1 shows that the average user repeatability was between 7 and 8.5%. On the other hand, for a given cluster, the repeatability was not the same for all users.

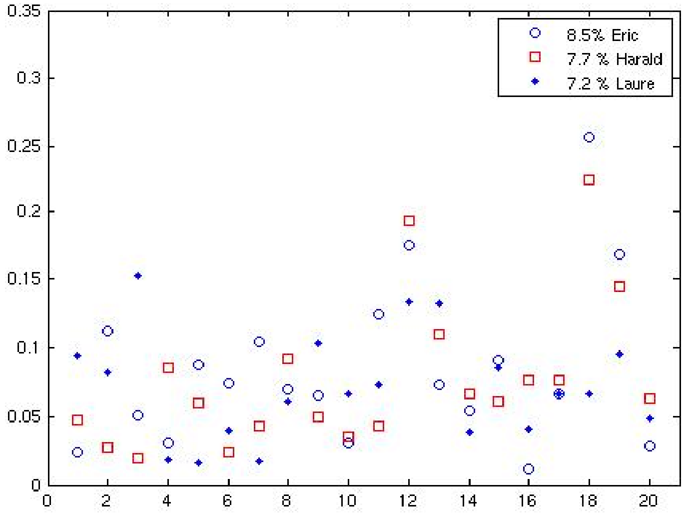

Figure 1. Repeatability of the three users for each cluster.

The data analysis shows that the level of repeatability was independent from cluster symmetry and density.

Reproducibility:

Figure 2 shows the reproducibility of each cluster, classified into 4 categories.

The average reproducibility per category increased slightly according to cluster asymmetry and density, but this criterion was not absolutely significant with this small number of clusters.

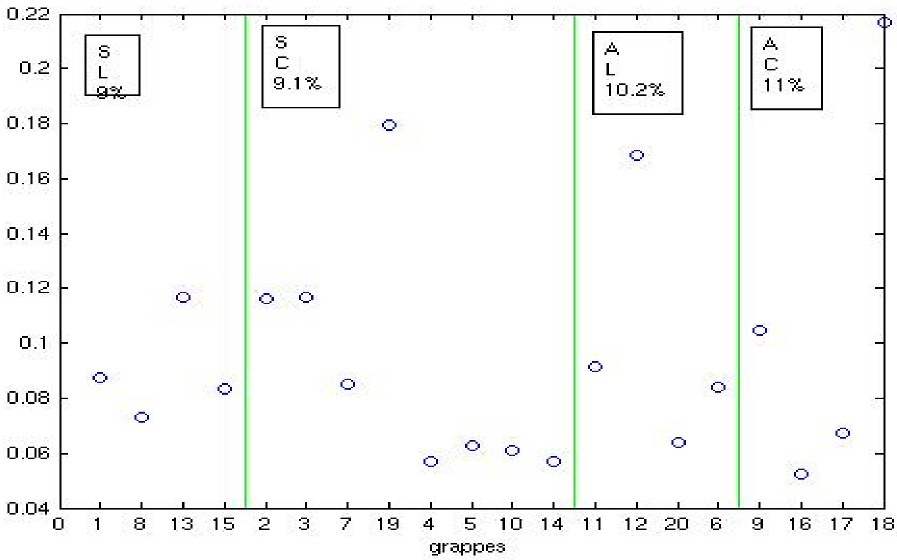

Figure 2. Reproducibility (M1) according to cluster categories

These first results showed that the user repeatabilities were satisfying, with coefficients of variation below 10% (7 –8.5%).

The reproducibility was equally around 10% with slight variations according to the type of clusters (symmetrical or not, dense or not). That is, they ranged from 9% for the easily measured clusters (symmetrical, loose) to 11% for winged and compacted clusters.

The difference of repeatability and reproducibility between cluster categories was not significant.

Correlation models for the estimation of average cluster weight

Regardless of the phenological stage, the correlation coefficient calculated by calibration for a given grape variety, vintage and phenological stage, was high. The averages obtained for all the grape varieties and vintages are presented in Table 1.

Table 1

Average correlation coefficients (r^2) for the linear relation between the cluster volume at different phenological stages and its weight at harvest (19 grape varieties, 4 vintages)

Stage	r^2 (calibration)	r^2 (cross-validation)
Setting	0.91	0.42
Bunch closure	0.90	0.71
Veraison	0.91	0.76

Using cross validation, the model performance was more heterogeneous but remained highly satisfying for the later growth stages.

At bunch setting, the errors were between 10 and 40 g (SECV) and thus, less acceptable.

Overall, regardless of the vintage or grape variety, the quality of the estimation was more precise if the measurement was carried out late. More specifically, there was an important difference between the results obtained at bunch setting and the ones obtained at the two other phenological stages (Figure 2).

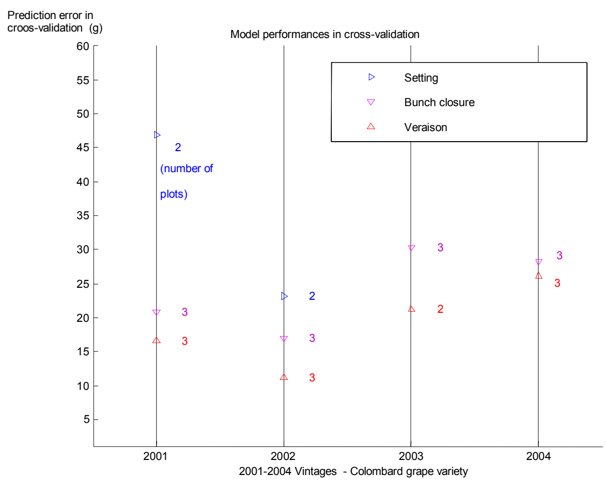

Figure 2

Model performance using Cross-validation with Colombard according to the phenological stage:

Figure 3 shows the coefficients of variation obtained with Cross-validation for all the grape varieties at bunch closure. The performance of the models present average errors of 10%, most of them being below 20%.

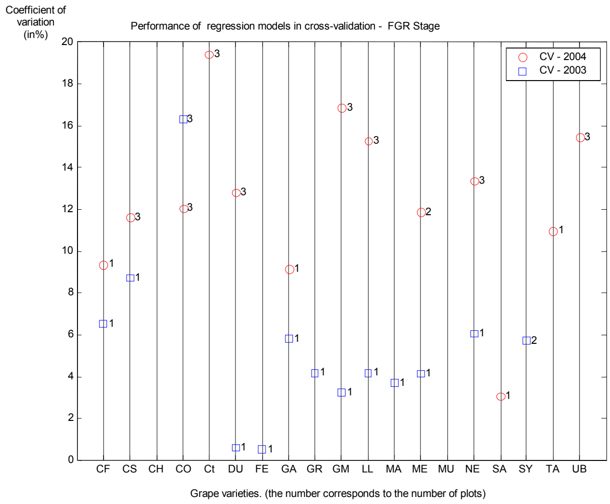

Figure 3

Coefficients of variation with Cross-validation of 19 grape varieties at bunch closure in 2003 and 2004

If the model was built using several plots per grape variety, the estimation performance decreased regardless of the phenological stage.

The same trends were observed for the data obtained at veraison. The coefficients of variation for the estimation of the average cluster weight per plot were slightly better (always below 13%).

In 2005, the development of models based on measurements performed with the REV allowed to obtain better results with more stable coefficients of variation and values below 10% (Table 2).

Table 2

Coefficients of variation with Cross-validation of REV measurements at bunch closure in 2005.

Plot	Grape variety	Real harvest weight (g)	Weight Estimated by REV (g)	Error (g)	Error(%)
1	Negrette	221	193	28	13%
2	Negrette	251	239	12	5%
3	Negrette	180	166	14	8%
1	Duras	262	243	19	7%
2	Duras	218	225	7	3%
1	Cabernet Sauvignon	111	112	1	1%
2	Cabernet Sauvignon	96	100	4	4%
3	Cabernet Sauvignon	100	92	8	8%
1	Cabernet Franc	111	102	9	8%
2	Cabernet Franc	135	131	4	3%
3	Cabernet Franc	111	101	10	9%

The utilisation of an optical sensor allowed to improve the precision of the models.

Validation of the models by estimating the average weight of the following year:

Among the 19 grape varieties studied, only 4 were measured over 4 years (from 2001 to 2004) at the phenological stages “bunch closure” and “veraison”: Colombard, Duras, Merlot and Negrette. A first analysis of the models showed that the latter were different according to the vintage (Figure 4). The 2003 vintage was atypical for 3 out of the 4 grape varieties studied (Colombard was the exception).

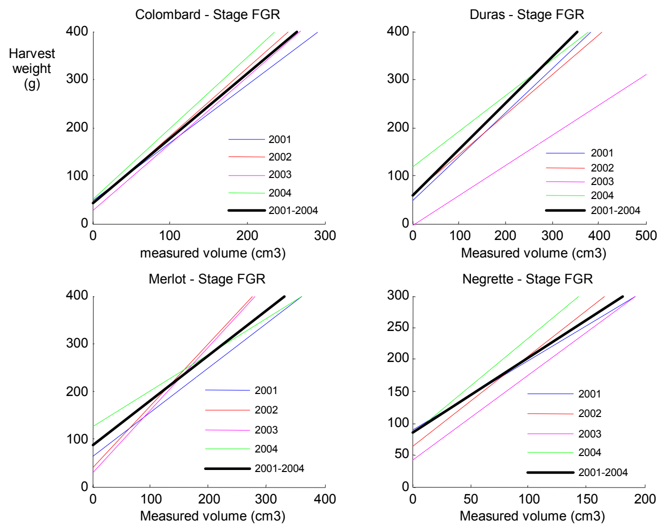

Figure 4

Correlation models of 4 grape varieties at bunch closure

Figure 4

Correlation models for the vintages 2001-2004 for 4 varieties at stage 2

The study focussed on the 2004 vintage using models that were built on all the possible combinations, which had been obtained in previous years (2001 to 2003).

If the models were built using data from 3 years, the coefficients of variation increased with average factors of 1.5 to 2 compared with the best performances obtained with the Cross-validation model for the same vintage.

Figure 5 shows the results of the estimations for the grape variety Negrette at bunch closure.

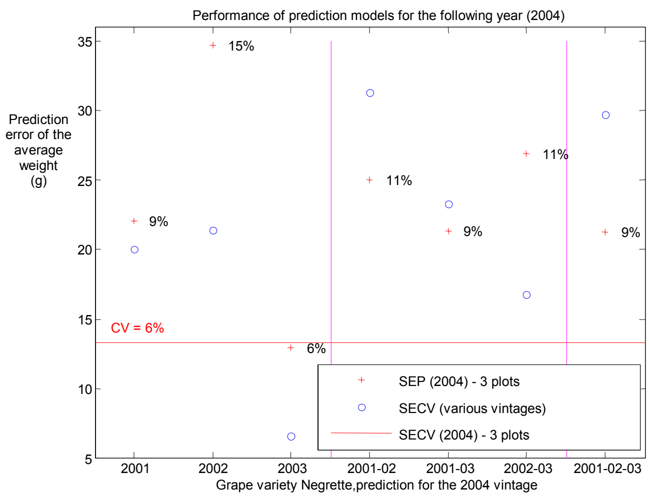

Figure 5

Model performance for the prediction of the 2004 vintage based on previous data

When the model was calculated with one year only, the performances differed significantly from year to year depending on the similarity between the model year and the year to be predicted.

If several years were considered for the model, the prediction stability increased, allowing to obtain a remarkable average performance of close to 9%.

CONCLUSIONS

This statistical study showed the relevance of measuring the cluster volume during the pre-veraison phase in order to estimate its weight at harvest.

Stable models could be demonstrated for the correlation between the measurement of the volume at bunch closure or veraison, and its weight at harvest.

When significant data was available for grape varieties, estimation errors remained below 10%. A plot effect could be demonstrated, but it was decreased by pooling the data obtained from several vintages.

The development of a portable optical sensor, the REV, allows a simple and reliable field application. Its utilization improves the precision of the prediction models.

BIBLIOGRAPHY

BOOYSEN J.H., ORFFER C.J., BEUKMAN E.T., 1978. Crop forecasting for vine grapes in South-Africa. *OVRI Stellenbosch, 78-94*.

GERBIER N., REMOIS P., 1977. Influence du climat sur la qualité et la production du vin de Champagne. *Monographie n°106, Météorologie Nationale*.

HUGLIN P., BALTHAZARD J., 1975. Variabilité et fluctuation de la composition des inflorescences et des grappes chez quelques variétés de V. vinifera. *Vitis, 6-13*.

MURISIER F., JEANGROS B., AERNY J., 1986. Maitrise du rendement et maturité du raisin. *Revue Suisse Vitic. Arboric. Hortic. Vol 18 (3) : 149-156.*

SCHNEIDER C., 1995. La prévision, un outil pour la maîtrise des fluctuations de rendement en viticulture. *C.R. GESCO, 240-246.*

SERRANO E., DUFOURCQ T., CHABERT M., 2002. Recherche d'une méthode simple et fiable d'estimation des rendements à la parcelle. *Actes de colloque Journée Technique Maîtrise des rendements en viticulture, dec. 2002. Station Régionale ITV Midi-Pyrénées.*

WURGLER W., LEYVRAZ H. ET BOLAY A., 1955. Peut-on prévoir le rendement de la vigne avant le débourrement ? *Annuaire agr. D. I. Suisse, 766-783*.