Estimating the seed content of *Vitis vinifera* cv. Gewürztraminer berries by two parameters related with the fecundation process

by

E. Duchêne[1], C. Schneider[1], C. Hennequet[2] and S. Huet[2]

[1] Unité de Recherches Vigne et Vin, INRA, Colmar, France
[2] Laboratoire de Biométrie, INRA, Jouy-en-Josas, France

S u m m a r y : The presence of a developed endosperm in a grape berry has been suggested to be a sufficient condition for berry set. The development of both, endosperm and embryo, leads to the formation of a seed and a grapevine flower usually contains 4 ovules. Taking this into account, a probabilistic approach has been used to study the seed content of Gewürztraminer berries in different plots in Alsace (France) over 7 years. The estimation of two probabilities, in relation with the development of the endosperm and the embryo, gives a good estimation of 5 variables (percentage of berries with 0 to 4 seeds). Knowing the number of flowers per shoot, this model explains 98 % of the variability of the number of berries per shoot and 99 % of the variability of the number of seeds per shoot between the years.

K e y   w o r d s : Vitis vinifera, seed, berry, flower, fruit set, fecundation, model.

Introduction

The grape berry weight has been shown to depend on the number of seeds (*e.g.* Olmo 1946; Huglin and Balthazard 1961; Gärtel 1961). Thus, the potential weight of berries per shoot mainly depends on the number of berries of this shoot, which is fixed after fruit set, and on the number of seeds per berry.

To better understand the effects of genotype, environment and cultural practices on the formation of seeds and berries, a precise definition of the main phenomena and of the variables which characterise them appears to be necessary.

Research on seed abortion (Pearson 1932; Olmo 1946; Barritt 1970) has shown, except for parthenocarpic varieties of the Corinth type, that if a berry set there is always a certain development of the endosperm in at least one seed. For wine grapes, Kassemeyer and Staudt (1982) have shown that in Gewürztraminer and Weisser Burgunder (Pinot blanc) shedding or limited growth of pistils occurred if all 4 ovules were aborted. Shedding of fertilized flowers was unusual, accounting only for 3-6 % of the total losses; in this case the endosperm had a maximum of 8 nuclei. With other varieties, Vallania *et al.* (1987) came to similar conclusions. With regard to small juicy berries, Staudt and Kassemeyer (1984) have suggested that for some varieties, including Gewürztraminer, there is always an endosperm in the largest seed of the berries; at least, from the size of these seeds it was concluded that fecundation had taken place. This is consistent with observations on seedless varieties. Previously, Kassemeyer and Staudt (1983), following other authors (review: Pratt 1971), had shown that the embryo development starts after the endosperm development and that this happens only if the endosperm has become cellular, *i.e.* about 20-21 d after flowering.

Based on these data, we have developed a model considering that the endosperm and embryo development process could be described by a probabilistic approach. This paper explains the principles of the model, the estimation of its parameters and demonstrates how to provide a simple description of the seed content of grape berries. The relationships between the parameters of the model and fruit set will then be discussed. The results are based on data of the variety Gewürztraminer.

Material and Methods

The model is based on the following hypotheses:
- A grapevine flower contains 4 ovules.
- The future development of each ovule is independent from that of the others.
- If at least one endosperm develops in a berry, this berry will reach maturity. On the contrary, if there is no fecundation at all in a berry, the berry will never develop and it will drop or stay green and small.
- The presence of both endosperm and embryo is necessary to get a normal seed.
  If we denominate
- p_e the probability that a fecundation process takes place and that an endosperm develops,
- p_s the probability that the fecundation leads to the formation of an embryo and furthermore a seed, knowing the endosperm is developed,
- b_4 the probability that a berry contains 4 normal seeds,
- b_3 the probability that a berry contains 3 normal seeds,
- b_2 the probability that a berry contains 2 normal seeds,
- b_1 the probability that a berry contains 1 normal seed,
- b_0 the probability that a berry contains no normal seed but at least one endosperm,

- b'_0 the probability that an ovary contains no endosperm, we can assume that Y, the random variable describing the seed content of a reproductive organ follows a multinomial law whose parameters are N, b_4, b_3, b_2, b_1, b_0, b'_0, where

- N is the total number of flowers,

- $b_4 = (p_e x p_s)^4$,

- $b_3 = 4x(p_e x p_s)^3 x(1-p_e x p_s)$,

- $b_2 = 6x(p_e x p_s)^2 x(1-p_e x p_s)^2$,

- $b_1 = 4x(p_e x p_s)x(1-p_e x p_s)^3$,

- $b_0 = (1-p_e x p_s)^4-(1-p_e)^4$,

- $b'_0 = (1-p_e)^4$.

Under our hypotheses, the probability for a flower not to develop is $b'_0 = (1- p_e)^4$. We can define a fruit set factor (FSF) as $FSF = 1- b'_0 = 1- (1- p_e)^4$ and the average number of seeds per berry (SPB) as $SPB = 4xp_e x p_s/(1- (1- p_e)^4)$.

If we consider only the number of set berries N_v, we can assume that the number of seeds per berry is distributed as a multinomial law Y_v with parameters N_v, b_{0v}, b_{1v}, b_{2v}, b_{3v}, and b_{4v} where $b_{iv} = b_i/FSF$ and i varies from 0 to 4.

In order to test this model, we need data describing the number of berries with 0, 1, 2, 3 and 4 seeds (N_{iv}, i from 0 to 4) in a sample of set berries (N_v). For additional relations to the fruit set ratio (FSR), the total number of set berries and the corresponding number of flowers are required.

Since 1984, more than 200 vine plots in Alsace (France) were checked for regional yield forecasts (Schneider 1995). These plots, planted with 8 varieties, are representative for the whole area (14,000 ha). All vines are cane-pruned according to the traditional Alsatian training system (Schneider 1989). More precise measurements concerning number of flowers, berries and seeds were performed on a sub-sample of this population, on a smaller geographic area (6,700 ha) with a smaller number of varieties.

Since 1992, three Gewürztraminer plants at 17 plots were examined three times of a year:

- Between two weeks to 3 d before flowering, the inflorescences of one shoot per plant were photographed. This shoot was chosen at a node position previously randomly determined, within the proximal third of a cane of the first plant, the middle third of the second plant and within the distal third of the third plant. If unexpectedly a shoot was not fertile, the closest fertile one was chosen. The total number of flowers per inflorescence was then estimated with a linear model (Schneider 1992); the equation is: $F = 1.56 \times B + 1.06$, where F = estimated number of flowers per inflorescence and B = counted number of flower buds on the photo.

- Approximately 35 d after the onset of flowering, a few days after complete fruit set, the clusters of the identified shoots were cut and berries with a diameter exceeding 3 mm were counted. The seed content of berries of a sub-sample of approximately 100 berries per cluster was determined. Seeds shorter than one third of the length of a normal seed were discarded. At this stage the seed number was determined only since 1995. These data are referred to as "fruit set data".

- At harvest, about 100 berries per plant were sampled. The berries were counted and weighed and the individual seed number was determined. These data are referred to as "harvest data".

Our main data file for Gewürztraminer contains 340 lines (7 years and approximately 48 plants per year) with 12 vari-

ables at the shoot level: number of flowers per shoot, number of set berries per shoot, number of berries with 0, 1, 2, 3 and 4 normal seeds in a sample, 35 d after flowering (since 1995) and at harvest time (since 1992).

The p_e and p_s parameters were estimated by maximising a criteria Cn equivalent to the likelihood of the observations, leading to the estimators $\hat{p}_e$ and $\hat{p}_s$.

$$Cn(p_e,p_s) = \sum_{i=0}^{4} N_{iv} Log(b_{iv}(p_e,p_s)),$$

where N_{iv} is the number of berries of the sample containing i seeds. The confidence regions for (p_e,p_s) were calculated using the log-likelihood ratio S_L. We compute S_L on a grid of values (p_{el},p_{sm}) for the parameters p_e and p_s,

$$S_L(p_{el},p_{sm}) = -2x[Cn(p_{el},p_{sm}) - Cn(\hat{p}_e,\hat{p}_s)].$$

Under the null hypothesis $H_0=\{(p_e,p_s)=(p_{el},p_{sm})\}$, S_L is distributed as a χ^2 with two degrees of freedom. The region of acceptance with the asymptotic confidence level 1-α is the set

$$\{(p_{el},p_{sm}) \text{ satisfying } S_L(p_{el},p_{sm}) \leq \chi^2_{1-\alpha}(2)\}$$

(Kendall and Stuart 1967).

If $\alpha = 5\%$, $\chi^2_{1-\alpha}(2) = 5.99$.

Calculations were carried out with a program running under Splus (MathSoft Inc., 1988, 1996, Version 3.4). A χ^2 statistic was also calculated to quantify the divergence between estimated and measured values.

$$\chi^2 = \sum_{i=0}^{4} \frac{(N_{iv} - N_v \hat{b}_{iv})^2}{(N_v \hat{b}_{iv})} \quad \text{where } N_v = \sum_{i=0}^{4} N_{iv} \text{ and } \hat{b}_{iv}$$

the estimator of b_{iv} deduced from $\hat{p}_e$ and $\hat{p}_s$.

Results

Tabs. 1 and 2 present the two raw data sets obtained just after fruit set and at harvest, respectively. Fig. 1 shows the results of the calculations of the seed number obtained 35 d after flowering (fruit set data) in 1995, 1996, 1997 and 1998.

Table 1

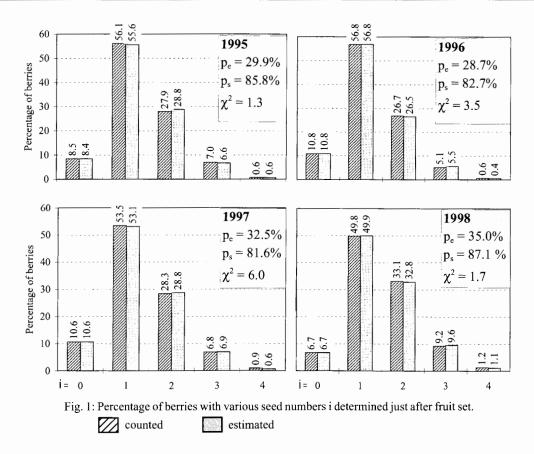Fruit set data: Seed number of berries collected just after fruit set. 17 (1997: 16) plots

Year	Number of plants per sample	Number of berries with					Total
		0 seed	1 seed	2 seeds	3 seeds	4 seeds	
1995	32[*]	220	1459	726	181	16	2602
1996	50	400	2104	991	189	22	3706
1997	45	657	3319	1757	420	53	6206
1998	47	362	2694	1790	500	64	5410

[*] Only 11 out of 17 plots were checked.

Table 2

Seed number per berry at harvest and the fruit set ratio determined on the same plants. 17 (1997: 16) plots

Year	Number of plants per sample	Number of berries with					Average number of			Fruit set ratio (FSR) (2)/(1) (%)
		0 seed	1 seed	2 seeds	3 seeds	4 seeds	seeds per berry	flowers per shoot (1)	berries per shoot (2)	
1992	50	161	2692	1887	527	46	1.55	371.9	202.2	54.4
1993	50	163	2977	2250	574	75	1.57	272.2	154.3	56.7
1994	47	170	2826	1240	194	10	1.34	279.2	134.2	48.1
1995	50	257	2570	1412	262	30	1.39	251.8	124.9	49.6
1996	50	526	2986	1372	251	17	1.27	140.4	77.2	55.0
1997	45	360	2440	1401	322	36	1.39	269.8	152.4	56.5
1998	48	151	2140	1839	469	54	1.60	189.8	122.2	64.4

Fig. 1: Percentage of berries with various seed numbers i determined just after fruit set.
counted   estimated

The greatest difference between an estimated $\hat{b}_{iv}$ and a measured proportion (N_{iv}/N_v) is 0.9 % and the maximum χ^2 is 6.0 in 1997. For all these data, the model represents adequately reality. The probability to get an endosperm (p_e) ranges from 29 to 35 % over the years. However, in 13-18 % ($1-p_s$) of the cases, this process does not lead to the formation of a seed.

In order to test the relationships between p_e, p_s and the actual fruit set with the maximum available data, we applied the model to the data acquired at harvest time. In this case, we can work with 7 instead of 4 years.

Again, the fittings between measured and estimated proportions are quite satisfactory (Fig. 2).

However, in comparison with the fruit set data, we can state that (i) appreciated with a χ^2 statistic, the quality of the fittings is poorer, especially in 1998 (Tab. 3), (ii) if there is a good correspondence between p_e estimated at fruit set or at harvest, values of p_s are always higher at harvest (Fig. 1 and Tab. 3). This may be explained by the fact that the proportion of 0-seeded berries is generally lower and the proportion of 2-seeded berries generally higher at harvest than after fruit set (Fig. 3), which may be due to the sampling procedure. After fruit set, measurements were based on entire clusters whereas at harvest measurements were based on small parts of bunches (5-10 berries). We can not exclude that large 2-seeded berries may have been chosen more often than small berries without seeds.
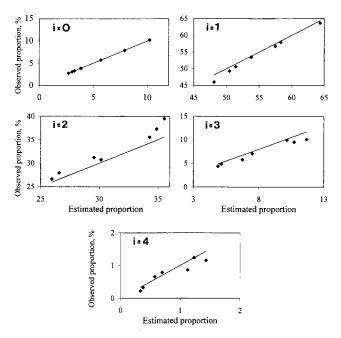
Fig. 2: Relationships between estimated and observed proportions of berries with various seed numbers i (harvest data). The line represents the bisectrix.

Table 3

Estimation of p_e, p_s, fruit set factor (FSF) and number of seeds per berry (SPB) determined from the harvest data

Year	p_c (%)	p_s (%)	χ^2	FSF (%)	SPB
1992	32.7	94.2	6.5	79.4	1.55
1993	33.3	94.7	19.7	80.2	1.57
1994	22.9	94.3	7.4	64.7	1.34
1995	28.2	90.6	11.3	73.3	1.39
1996	27.2	84.0	2.1	72.0	1.27
1997	31.6	86.2	2.3	78.1	1.39
1998	35.3	93.4	37.7	82.5	1.60

The confidence regions at the 5 % level for the estimations of p_e and p_s at harvest are represented in Fig. 4; it clearly shows that the variations of p_e and p_s are independent. Indeed, it seems possible to observe a low p_e together with a high p_s (1994), as well as a high p_e and a low p_s (1997).

From 1995 to 1998, the p_e parameter is similar for both, harvest and fruit set data (Fig. 1 and Tab. 3); thus we assume that the harvest data set gives a good estimation of this parameter for the years 1992 to 1994. The probability for a berry not to drop and to become juicy $\left(\hat{FSF} = 1 - (1 - \hat{p}_e)^4\right)$ has been calculated for each year using the harvest data set (Tab. 3). These values are high compared with the actual FSR (Tab. 2) and the corrrelation with FSR is not as good as expected ($R^2 = 0.68$). For example, although the estimations of FSF are similar for 1995 and 1996, the actual FSR values are quite different. The same statement can be made for 1993 and 1998. A theoretical number of berries per shoot can be calculated by multiplying the number of
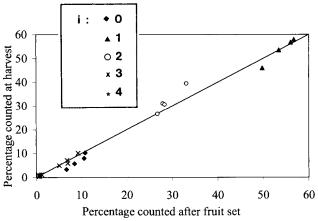

Fig. 3: Relationship between seed number i of berries counted after fruit-set and at harvest.
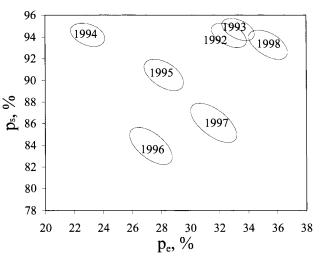

Fig. 4: Confidence regions for parameters estimated from harvest data (5 % level).

flowers per shoot by the corresponding probability for a flower not to drop. Fig. 5 shows (i) that the actual number of berries per shoot is always lower than the theoretical number, (ii) that the difference between the estimated and the actual number of berries per shoot seems to be linked to the number of flowers per shoot ($R^2 = 0.89$). The higher the number of flowers per shoot, the greater the difference.

Finally, if we take into account the number of flowers per shoot and $\hat{p}_e$, we can explain 98 % of the variability of the number of berries per shoot. If we add the $\hat{p}_s$ parameter, we can calculate the average seed number per berry $(\hat{SPB})$ [a] and a theoretical seed number per shoot. In this case, we can explain 99 % of the variability of the seed number per shoot.

Discussion

The proposed model represents seed distribution in berries with good accuracy. With two parameters, based on biological considerations, 5 variables (percentage of berries with 0, 1, 2, 3 and 4 seeds) can be estimated. However, the
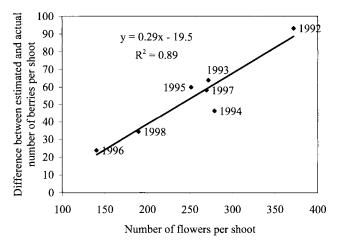
Fig. 5: Relationship between the number of flowers per shoot and the difference between the calculated and the measured number of berries per shoot.

sampling method is a critical point. To get reliable data of the distribution of seeds, the analysis of a few entire clusters is preferred to picking small bunches from a lot of clusters. We can put forward another argument to test the validity of the model. Considering that the model is valid and the estimations of the parameters p_e and p_s are true values, the calculated χ^2 between predicted and measured values would follow a χ^2 law with two degrees of freedom (5 classes and two estimated parameters). In this case, we would accept the null hypothesis (the number of berries counted in each category is compatible with the forecast), at the 5 % level for the 4 years with fruit set data.

The model is currently being tested for different varieties, for individual plots or single clusters and the first results are quite positive, except for a few varieties such as Riesling. For Gewürztraminer, we would need tests in situations with very low fruit set. The model offers new possibilities for analysing the effects of growth or climatic conditions on seed formation. Indeed, the p_e parameter is the result of what happens before and during flowering whereas p_s synthesises what happens between fecundation and the onset of embryo development, about 21 d after anthesis according to KASSEMEYER and STAUDT (1983).

The ability of this model to describe fruit set requires further attention. The actual number of set berries is always lower than the predicted one. We can put forward two hypotheses to explain this.

(1) The difference is due to young berries which drop. This does not correspond to the hypotheses of the model nor to literature, indicating that the maximum value is 6 % of the total losses (see Introduction). Moreover, loss of young berries would certainly affect berries with 0, 1, 2, 3 or 4 seeds in different ways. In this case, there are few chances that the distribution of seeds can be described with parameters related to fecundation. However, we have no definitive argument to exclude this hypothesis.

(2) Individual flowers have different abilities to become berries. We already have taken this idea into account. As a matter of fact, the p_e parameter includes the probabilities:
- for an ovule to exist and to be functional,
- for a pollen tube to reach an ovule,

- for the fecundation process to take place in good conditions,
- for the endosperm to reach a certain developmental stage.

In the category of non-set organs, associated with the $(1-p_e)^4$ probability, we already include flowers with no functional ovules which can never develop to berries.

To be consistent with our results, we have to make the hypothesis that a population of flowers does not follow our model and is never able to produce berries. This population corresponds to approximately one third of the total number of flowers (Fig. 5). This idea has already been suggested by other authors (MULLINS et al. 1992). MAY (1987) has shown the wide range of flower sizes in one inflorescence. According to his data, a 1:3 ratio in width can be observed on flowers within a branch of a bunch at a given date. The size of flowers seems to depend on their position at the branch, flowers in terminal positions ("king" flowers) being larger than the others. Moreover, he observed, at least in one year, that flowers in terminal positions had a higher frequency to set than the others. EBADI et al. (1995) also noted that ovules of "king" flowers were bigger than those in "central-lateral" and "lateral-lateral" flowers. These results support the idea of heterogeneity within the flower population and that flowers have different probabilities to develop to berries. Further research is required to confirm these observations and the relations between "functional flowers", the probability to get an endosperm (p_e) and actual fruit set.

Acknowledgements

References

BARRITT, B. H.; 1970: Ovule development in seeded and seedless grapes. Vitis 9, 7-14.

EBADI, A.; MAY, P.; SEDGLEY, M.; COOMBE, B. G.; 1995: Effect of low temperature near flowering time on ovule development and pollen tube growth in the grapevine (Vitis vinifera L.), cvs Chardonnay and Shiraz. Aust. J. Grape Wine Res. 1, 11-18.

GÄRTEL, W.; 1961: Einfluss der im Weinbau gebräuchlichen Planzenschutzmittel auf Keimung und Schlauchwachstum bei Rebpollen. Mitt. Biol. Bundesanstalt für Land- und Fortswirtschaft 33, 108-112.

HUGLIN, P.; BALTHAZARD, J.; 1961: Observations préliminaires concernant la coulure du Muscat Ottonel. Vins d'Alsace 57, 122-129.

KASSEMEYER, H. H.; STAUDT, G.; 1982: Cytologische Untersuchungen über die Ursachen des Verrieselns bei Vitis. Vitis 21, 121-135.

- -; - -;1983: Uber das Wachstum von Endosperm, Embryo und Samenanlagen von Vitis vinifera. Vitis 22, 109-119.

KENDALL, M. G.; STUART, A.; 1967: The Advanced Theory of Statistics - Vol. 2. Charles Griffin and Co Ltd. London.

MAY, P.; 1987: The grapevine as a perennial, plastic and productive plant. Proceedings 6th Australian Wine Industry Technical Conference, 40-49. Australian Industrial Publishers, Adelaide (South Australia).

MULLINS, M. G.; BOUQUET, A.; WILLIAMS, L. E.; 1992: Biology of the Grapevine. Fruit Growth and its Regulation. Cambridge University Press, Cambridge.

OLMO, H. P.; 1946: Correlations between seed and berry development in some seeded varieties of *Vitis vinifera*. Proc. Amer. Soc. Horticult. Sci. **48**, 291-297.

PEARSON, H. M.; 1932: Parthenocarpy and seed abortion in *Vitis vinifera*. Proc. Amer. Soc. Horticult. Sci. **29**, 169-175.

PRATT, C.; 1971: Reproductive anatomy in cultivated grapes - a review. Amer. J. Enol. Viticult. **22**, 92- 109.

SCHNEIDER, C.; 1995: La prévision, un outil pour la maîtrise des fluctuations de rendement en viticulture. 8è journées du Groupement Européen d'Etudes des Systèmes de Conduite de la Vigne (GESCO), 3-5 Juillet 1995, Vairao/Portugal, 240-246.

- -; 1992: Analyse des composants de la production viticole alsacienne. Recherche d'une méthode fiable de mesure du nombre de fleurs/inflorescence. C. R. 4è Symposium International sur la Physio-logie de la Vigne, Turin/Italie, 11-15 Mai 1992, 133-136.

- -; 1989: Etude comparative de plusieurs systèmes de conduite sur deux cépages d'Alsace. Bull. OIV **62** (705-706), 781-793.

STAUDT, G.; KASSEMEYER, H. H.; 1984: Entstehen kleine Beeren bei *Vitis vinifera* durch Parthenocarpie? Vitis **23**, 205-213.

VALLANIA, R.; BOTTA, R.; ME, G.; 1987: Investigations on anomalies of ovule development and on pollination in mutated grapevines cv. Barbera. Vitis **26**, 1-8.