IMPROVED VINEYARD SAMPLING EFFICIENCY USING AERIAL NDVI

Authors:Jim MEYERS[1], Nick DOKOOZLIAN[2], Casey RYAN[2], Cella BIONI[2], Justine VANDEN HEUVEL[1]

[1]*Horticulture Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853 and Geneva, NY 14456*

[2]*Viticulture, Chemistry and Enology, E&J Gallo Winery, 600 Yosemite Blvd., Modesto, CA 95354*

*Corresponding author: *Justine@Cornell.edu*

Abstract:

Context and purpose of the study - Random sampling is often considered to be the best protocol for fruit sampling because it is assumed to produce a sample that best represents the vineyard population. However, the time and effort in collecting and processing large random samples can be cost prohibitive. When information about known field variability is available, a spatially-explicit sampling protocol can use that information to more efficiently sample the vineyard population. A commonly used method for mapping vineyards is normalized difference vegetation index (NDVI) which can be acquired through satellite imagery or overhead flight by plane or drone. This study seeks to improve sampling efficiency by using aerial NDVI vineyard imagery to compute optimal spatially-explicit sampling protocols that minimize both the number of locations sampled and the time required to sample, while also minimizing potential of human errors during data collection.

Material and methods - NDVI imagery acquired from LANDSAT 7 was used to map spatial variability, at a resolution of 30 by 30 meter pixels, in 24 vineyards located in California's Central Valley. Three sampling methods, each sampling twenty whole fruit clusters, were compared to determine relative efficacy: 1) Twenty pixels selected by a random number generator (RAND20); 2) Four fixed locations, representing each quadrant, near the edge of the vineyard sampling two pixels at each location (RAND4x2), and; 3) One location, determined by a novel optimization algorithm, sampling three pixels (NDVI3). The vineyards were sampled weekly between verasion and harvest to measure Brix, titratable acidity (TA), pH, and total anthocyanins.

Results – All three sampling methods were highly correlated in pair-wise comparisons of Brix (R= 0.86 – 0.93), TA (R= 0.93 – 0.96), pH (R= 0.96 – 0.98), and anthocyanins (0.88 – 0.90). Comparing NDVI3 and RAND4x2 to RAND20, deviation from RAND20 measurements was slightly lower in NDVI3 for Brix, TA, and pH, and slightly higher for anthocyanins. These results suggest that vineyard sampling in a single row and an optimally calculated location can produce results similar to more costly random sampling.

Keywords: Grapevine, Sampling, NDVI, Optimization, Spatial variability, Efficiency.

1. Introduction

Spatial patterns in agricultural plots were acknowledged well before they could be imaged (Student 1938, Jefferys 1939), resulting in the use of randomized complete block designs (RCBD) to compensate.Vine sampling protocols have been modified by researchers to compensate for unknown spatial variability within the canopy (Roessler and Amerine 1958, Rankine et al. 1962, Iland et al. 2004) but only recently a protocol was proposed to directly sample a known measured population distribution in winegrape canopies (Meyers et al. 2011). Optimization of sampling operations using known spatial patterns can lead to a lower information cost (Zilliak 2014). Limiting sampling locations to a single vineyard row can both reduce sample sizes and the labor required to collect the sample (Meyers and Vanden Heuvel, 2014). The sampling protocols described in Meyers and Vanden Heuvel (2014) demonstrated the use of 0.5-meter pixel-scale aerial NDVI imagery to compute an optimal sampling strategy by directing sampling to a set of spatially explicit pixel locations, but requiring a field technician to sample with such a high degree of spatial precision has practical limitations. Therefore, the objective of this work was to determine if sampling efficiency could be improved by reducing the total number of sampling locations and the distance between them.

2. Material and methods

Aerial NDVI Imagery. All NDVI imagery was LANDSAT 7 30-meter scale images (each pixel covers 900 square meters of land) captured at early veraison between 2012 and 2016. Experimental vineyard blocks were bounded by shapefiles created using ArcMap (Version 10.5, Esri, Redlands, CA) which were manually processed to eliminate features such as access roads, buildings, and intra-block trees from the block image pixels.
Vineyards. All experimental vineyards were in California's Central Valley. Thirteen blocks were used In 2016, 24 blocks in 2017.
Sampling Protocols. Three sampling protocols were employed in the study (Table 1):
RAND20. Twenty locations were randomly chosen via custom software (MATLAB version 2017a, MathWorks, Natick, MA). Locations were computed by selecting 20 random pixels from the block shapefile, with vineyard edges avoided.
RAND4x2. Four locations were chosen for each block, with two pixels sampled at each location with vineyard edges avoided.
NDVI3. The design goal of the NDVI3 protocol was to find the best location in a vineyard block that captures the population distribution of NDVI pixel values for the block within three consecutive pixels.
Comparison of Sampling Protocols. All statistics were performed using MATLAB (version 2017a, Natick, MA).Pearson Correlation Coefficients were used to compare fruit chemistry measurements between RAND20 and RAND4x2, and between RAND20 and NDVI3.Two-value Kolmogorov-Smirnov testing was used to compare sampled pixels to block population pixels (Stephens 1992).

3. Results and discussion

Correlations among sampling methods. Pearson Correlation Coefficients between RAND20 and RAND4x2, and between RAND20 and NDVI3 are presented in Table 2. With one exception, correlation coefficients were >0.85. In 2016 correlations were all high, suggesting that both RAND4x2 and NDVI3 were suitable alternatives to RAND20 for measuring Brix, TA, and total anthocyanins. The combined 2017 data suggests that NDVI3 is a better alternative to RAND20 than is RAND4x2, purely in terms of sampling accuracy, even without considering efficiency gains.
Sample fitness. For 12 of the 13 blocks, the K-S test p value was lower for RAND4x2 than for NDVI3 (higher p values = better fit). All NDVI3 solutions in 2017 passed the K-S test (data not shown).

4. Conclusions

The results suggest that the demonstrated NDVI3 sampling protocol is functionally equivalent to or better than both RAND20 and RAND4x2 in its ability to accurately estimate population Brix, TA, pH, and total anthocyanins. Because NDVI3 sampling requires that a field technician locate and visit only one location within a block, rather than four for RAND4x2 or twenty for RAND20, it requires less time to perform.

5. Literature cited

van Es, H.M., Gomes, C.P., Sellmann, M. and van Es, C.L. (2007) Spatially balanced complete block designs for field experiments. Geoderma 140:346–352.

Jeffreys, H. 1939. Random and systematic arrangements. Biometrika 31:1-8.

Meyers, J.M., Vanden Heuvel, J.E. 2014. Use of normalized difference vegetation index images to optimize vineyard sampling protocols. American Journal of Enology and Viticulture 65(2) 250-253.

Meyers, J. M., Sacks, G. L., van Es, H. M., & Vanden Heuvel, J. E. 2011. Improving vineyard sampling efficiency via dynamic spatially-explicit optimisation. Australian Journal of Grape and Wine Research, 17, 306–315.

Panten, K. and R.G.V. Bramely. 2012. Whole-of-block experimentation for evaluating a change to canopy management intended to enhance wine quality. Australian Journal of Grape and Wine Research 18:147-157.

Rankine, B.C., Cellier, K.M. and Boehm, E.W. 1962. Studies on grape variability and field sampling. American Journal of Enology and Viticulture 13:58–72.

Student. 1938. Comparison between balanced and random arrangements of field plots. Biometrika 29:363–379.

Stephens M.A. .1992. Introduction to Kolmogorov (1933) On the Empirical Determination of a Distribution. In: Kotz S., Johnson N.L. (eds) Breakthroughs in Statistics. Springer Series in Statistics (Perspectives in Statistics). Springer, New York, NY.

Table 1. Overview of sampling methods used in the study. Number of sampling locations refers to the number of predefined locations, expressed as GPS coordinates, are sampled in each block. Number of pixels per location refers to the number of 30 x 30 meter Landsat pixels traversed at each location during sampling. Pixels sampled is the product of locations and pixels per location. Number of clusters sampled totaled 20 per block for all methods.

Treatment	Description	# sampling Locations	# Pixels / Location	# Pixels Sampled	# Clusters Sampled
R20	Computer generated random pixel slection	20	1	20	20
CM8	Four vineyards rows, one in each quadrant, sampled near end of row	4	2	8	20
NDVI3	Three adjacent pixels representing the three quantiles of population	1	3	3	20

Table 2. Pearson correlation coefficients comparing random sampling (RAND20), sampling using a three-pixel NDVI directed protocol (NDVI3), and sampling using an eight-pixel quadrant directed protocol (RAND4x2) for Brix, titratable acidity (TA), pH, and total anthocyanin concentration (Anthos). RAND20 was performed once in 2016 (just after veraison onset) and twice in 2017 (just after veraison onset and again close to harvest).
N = number of blocks sampled. ND = not determined due to lack of measurement.

		Brix		TA		pH		Anthos	
	R20 Timing	NDVI3	CM8	NDVI3	CM8	NDVI3	CM8	NDVI3	CM8
2016	First (N=13)	0.85	0.89	0.91	0.93	ND	ND	0.98	0.91
2017	First (N=16)	0.91	0.89	0.89	0.85	0.96	0.93	0.83	0.94
	Second (N=16)	0.73	-0.05	0.90	0.78	0.98	0.95	0.95	0.88
	Combined (N=32)	0.93	0.86	0.96	0.93	0.98	0.96	0.90	0.90