Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-16 (2016) 398–403

Non-Productive Vine Canopy Estimation through Proximal and Remote Sensing ⋆

Julie Tang * **Michael Woods** ** **Stephen Cossell** * **Scarlett Liu** * **Mark Whitty** *

* *University of New South Wales, Australia (e-mail: {julie.tang; s.cossell; sisi.liu; m.whitty}@unsw.edu.au).*
** *e-mail: m.woods1987@gmail.com*

Abstract: Non-productive canopy detection in a viticultural block is a key factor in reducing the drain on infrastructure and improving management practices. However, current methods are significant in cost, biased, and do not provide information on location of non-productive canopy. This paper proposes both a proximal and remote sensing method for assisting in decision support and yield estimation from available technologies. The proximal method utilizes two different measures of green pixel thresholding in video frames, with results providing a useful relative measure of productivity across a vineyard block at the phenological stage of shoots. The remote sensing method utilizes local thresholding and Self-Organizing-Maps on aerial imagery to identify missing vines and total non-productive canopy on a block level. Results indicate the success of this semi-supervised method in providing a useful measure of non-productive canopy at the phenological stage of veraison; laying the groundwork for improved methods in this area. These methods provide practical outputs that lay the foundations for improving management decisions in an automatic and low-cost manner at different times in the season.

Keywords: NDVI, yield estimation, decision support systems, automation, mapping, machine learning

1. INTRODUCTION

An area that is of great importance, but has been not at the forefront of research is in classification of productive and non-productive land; using both proximal as well as aerial imagery. Non-productive land represents a drain on infrastructure and inputs such as irrigation and spraying. Although research in grapevines through spectral indices, photogrammetry, as well as proximal sensors such as Laser Range Finders (LiDAR) is comprehensive; much of the research has been in estimating and improving vine vigour Mathews and Jensen (2013); Llorens et al. (2011), fruit yield Nuske et al. (2014), and maturity Hall et al. (2011); Serrano et al. (2012).

The majority of existing methods for vine row extraction has been used in identifying variation in measures of vine health through Normalized Difference Vegetation Index (NDVI), Plant Cell Density (PCD), or Leaf Area Index (LAI) and obtaining relationships with the above mentioned research areas. The techniques previously used include spectral band thresholding Hall et al. (2003), skeletonization Nolan et al. (2015), and variations of Hough lines Comba et al. (2015).

Grapevines represent a high value crop and by providing information about productive and non-productive land within a vineyard to the farmer this can allow for the land to be better managed, so as to optimise the land use within each vineyard. With this information the farmer can make informed management decisions on vine replanting, however, current methods are costly and do not provide accurate locations of non-productive canopy within a vineyard.

The current methods for estimating productive and non-productive canopy include a rough estimation by the farmer, as well as manual counting to a lesser extent. Estimation by farmers is anecdotally known to have significant errors with large variation, as the approximation is based on their knowledge of missing vines. Manual counting techniques can also be error prone if taken at a very early phenological stage which is otherwise referred to as Modified Eichhorn-Lorenz (E-L) Coombe (1995) stage. This occurs when canopy has not yet fully developed. Thus, care needs to be taken in the timing of measurements as this can result in a large variation between actual non-productive canopy and estimated canopy extent of each vine.

In manual counting, regions with approximately 50cm of missing vines, dead vine arms, and "bare-wire" are recorded as a count to obtain a percentage of non-productive canopy per vine row or block. The estimation of non-productive canopy may be difficult if the vine canopy is not fully developed, especially early in the season. Currently, manual counting only reports the total distance or percentage of non-productive canopy per vine row or per block; reducing the manual counting effectiveness as localization of non-productive canopy from this data alone is difficult. Thus, manual measurements are often significant in cost and ineffectual. Therefore, an automated method is required to reduce the discrepancies in identification of the productive and non-productive vines with geo-location

⋆ This work was supported by Wine Australia

10.1016/j.ifacol.2016.10.073

mapping which greatly assists with vine block management.

Identification of non-productive areas on other broad-acre crops generally utilise a combination of multispectral, hyperspectral and thermal imagery such as Calderón et al. (2013); Turner et al. (2011) mounted on an Unmanned Aerial System (UAS). This allows for enough spatial resolution in the image to be obtained and thus an analysis on the non-productive areas.

This paper presents a novel method for sensing and classifying non-productive vine canopies from a combination of proximal camera data as well as multispectral aerial photography. The proposed aerial method firstly identifies missing vines and 'bare-wire' through thresholding techniques and then utilizes this information to further classify all non-productive canopy. The objective of the proximal method is to provide a comparison between ground-based and aerial techniques at an early stage in the season.

2. DATA COLLECTION

Aerial photography was obtained around veraison in 2015, across four different vineyard blocks and two trellis systems, situated in contrasting climatic regions. The site at Clare, South Australia contained one Chardonnay (40A) and one Shiraz (47A) block, both with sprawling canopies. The cool climate site at Orange, New South Wales, contained one Chardonnay (B12) and one Shiraz (B4) block, both with Vertically Shoot Positioned (VSP) canopies. The spatial resolution of the aerial imagery was 0.4m for the Clare vineyard blocks and 0.25m for the Orange vineyard blocks.

Proximal data for all vineyard blocks in the study was also captured at each phenologically significant stage; the data consists of vehicle mounted Go-Pro footage, which was collected in the same manner as prior work by Liu et al. (2015).

Manual counting of 50cm segments of non-productive canopy and number of missing vines per row were only available for the two Clare experimental blocks. These were taken at E-L stages 4 for Block 40A and 7 for Block 47A.

3. EXPERIMENTAL APPROACH

Using the approach of Hall et al. (2003), vines are segmented in an attempt to locate the edges of a vineyard block. Although such approaches are suitable for identifying existing rows, they are not suitable for obtaining the outline of a concave shaped block, as shown in Fig. 1. The idea of automatically determining row-end post locations from visible vines tends to under-represent the amount of non-productive canopy due to the possibility of missing vines at each row-end, as shown in Fig. 2. Thus, row-ends of each experimental block were marked with GPS locations and used for localizing rows.

3.1 Identifying missing vines from aerial imagery

Two methods were considered for segmenting vines. Supervised learning methods were found to perform well if

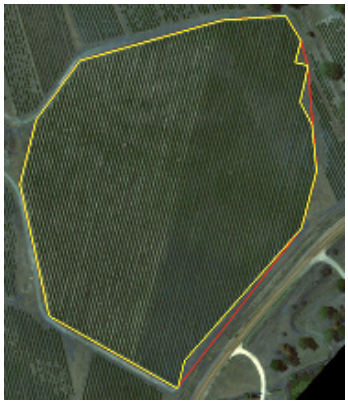

Fig. 1. Convex Hull (red) of a vineyard block which does not follow actual block outline (yellow)

Fig. 2. Missing vines at the row-end of a vineyard block results in misrepresentation of end-post location

at least 10 % of a block was labelled, however, this can't be considered automatic. Local thresholding was found to perform best, with a much easier implementation when used together with a large local neighbourhood. The local thresholding was performed by firstly re-sampling all aerial data to 0.1m spatial resolution using bi-linear sampling and then applying a tuned threshold parameter to segment the vines.

Similar to Nolan et al. (2015), we found a neighbourhood of vine row spacing, 3m in our case, was sufficient to prevent over-segmentation. However, the method we employed was that of Sauvola and Pietikäinen (2000) on the green spectral band as the histogram filtering method used by Nolan et al. (2015) required imagery of sufficiently high resolution, decimeter or less, which was not available for our study blocks. The histogram filtering method resulted in over-segmentation of thin or unhealthy vines which were present in one of our study blocks.

Although each vine-row end post was labelled, vines were not always situated along the straight line formed between the two opposite vine-row end posts. The vines that do exist on the straight line are herein referred to as row-centered pixels. Thus, in the case of straight rows, vines deviating from row-centered pixels need to be accounted for, this variation could be due to vine posts being damaged or the vine growing at different angles.

An indication of missing vines can be obtained by taking vine pixels within a 1 meter buffer on either side of the row-centered pixels, as shown in Fig. 3, on the green spectral band thresholded image. This can be found by taking the sum of the number of non-zero pixels in the thresholded

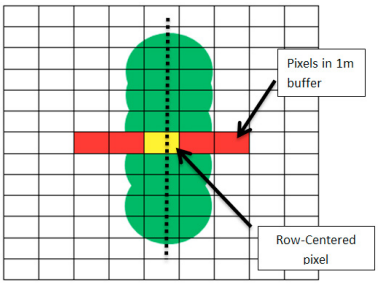

Fig. 3. A representation of pixels used to form a feature set (red) centered about a row-centered pixel (yellow). The pixels used are within 1m of the row-centered pixel on a line perpendicular to the vine-row center (black dotted line).

image. If more than one non-zero pixel exists, the row-centered pixel is assigned as vine.

The bare-wire pixels are thus found to be the areas where the sum of the pixels in the thresholded image equals 0. The length of bare-wire segments l_b from the image can then be calculated as the number of adjacent bare-wire pixels multiplied by a pixel to meter conversion, which is 0.1m/pixel. Missing vines were defined as a subset of homogeneous regions of linear "bare-wire". The number of missing vines N_{miss} is defined as in (1).

$$N_{miss} = \begin{cases} 0 & ; l_b < 0.5l_s \\ 1 & ; 0.5l_s < l_b \leq l_s \\ 2 & ; l_s < l_b \leq 2l_s \\ 2 + \lfloor \frac{l_b - 2C_b}{l_s} \rfloor & ; l_b > 2l_s \end{cases} \quad (1)$$

where l_s is the vine spacing length, the average vine spacing in a block is based on the average number of vines between evenly spaced posts, C_b is a buffer that is used to account for the possibility of overgrown canopies on both sides of the 'bare-wire' segment. The value of $C_b = 0.5l_s$ in this work has been selected as it provides the best representation of the growth pattern for the vines used in this study.

In the event where the missing vines are located on the end of rows, the number of missing vines that should be attributed to the section is $\lceil \mathrm{N}_{miss} + \frac{C_b}{l_s} \rceil$.

3.2 Identifying non-productive canopy

Identification of non-productive canopy can be non-trivial for low spatial resolution aerial imagery, such as that of block 40A and 47A. This causes vine cordons of thin or unhealthy vines to be represented as a mixture of cordon and ground-cover material; making visual classification difficult. Thus, a semi-supervised learning approach was taken to further classify the vine-row automatically.

Firstly, a number of features as defined in (2) - (7) are to be determined. A binarised mask is calculated using Sauvola thresholding on the Green channel, reducing the effects of the ground cover on the generation of vine features. The features chosen are calculated from the masked Red, Green, Blue, and Near-InfraRed channels from the acquired aerial imagery. Using the masked channels, all non-row pixels can be set to zero; row pixels maintain their original values. The features generated are ratios, which were chosen to reduce the effect of variations in lighting intensity throughout non-normalised aerial images. The features for the kth pixel are found using the following equations,

(1) Normalized Difference Vegetation Index

$$f(k)_1 = \frac{\phi_{NIR} - \phi_R}{\phi_{NIR} + \phi_R} \quad (2)$$

(2) NDVI-Blue

$$f(k)_2 = \frac{\phi_{NIR} - \phi_B}{\phi_{NIR} + \phi_B} \quad (3)$$

(3) NDVI-Green

$$f(k)_3 = \frac{\phi_{NIR} - \phi_G}{\phi_{NIR} + \phi_G} \quad (4)$$

(4) Green-Red Ratio

$$f(k)_4 = \frac{\phi_G - \phi_R}{\phi_G + \phi_R} \quad (5)$$

(5) Green-Blue Ratio

$$f(k)_5 = \frac{\phi_G - \phi_B}{\phi_G + \phi_B} \quad (6)$$

(6) Plant Cell Density

$$f(k)_6 = \frac{\phi_{NIR}}{\phi_R} \quad (7)$$

where ϕ_B, ϕ_G, ϕ_R and ϕ_{NIR} are the spectral reflectance measurements for the bands centered about 450nm, 550nm, 670nm and 780nm respectively.

The features were obtained for each pixel k, $F(k) = [f(k)_1, \ldots, f(k)_6]$ within the 1m radius of the row-centered pixels, as seen in Fig. 3. Features for the n pixels within the radius were considered as part of the same feature vector, $T = [F(1), \ldots, F(n)]$. As the row-centered pixels may not correspond to the center of a vine, the feature vector T was sorted in descending order of the $f(k)_1$ value for each pixel.

The classification method that is utilised in this paper is Self-Organizing-Maps (SOM) which was developed by Kohonen (1990) and is an unsupervised artificial neural network for automatic clustering and visualisation of data. Training occurred on a 4x4 SOM over 300 iterations, with 50% of 'bare-wire' classed samples, from Section 3.1 and 50% of random samples throughout the block. With an initial neighbourhood of 6 nodes, a unified distance matrix between nodes can be obtained, as shown in Fig. 4. Non-productive canopy nodes N_{np}, can then be obtained from the unified distance matrix through the following steps, where a dominant node N_D is a node that provides the most results for a classification.

The data from the block can then be classified into non-productive and productive canopy. In the case of no "bare-wire" detected in Section 3.1, the user can indicate a patch of "non-vine" ground-cover for training purposes.

In the proximal case, the use of a background at the phenological shoots stage allowed the number of green pixels to be assessed in the GoPro video imagery. Two methods were used to measure the proportion of green pixels within each frame. First, pixels from each frame were mapped to the HSV color space and partitioned either as green or non-green pixels using (8).

Algorithm 1 Identification of non-productive nodes

1: Find the $N_{D,1}$ of bare-wire training samples
2: Add $N_{D,1}$ into N_{np}
3: Remove all bare-wire instances that belong to $N_{D,1}$
4: Find the $N_{D,2}$ of the remaining bare-wire training samples
5: **if** size of bare-wire instances belonging to $N_{D,2} > 10\%$ of total training bare-wire samples **then**
6: include $N_{D,2}$ into N_{np}
7: **end if**
8: For the identified nodes, add additional closely associated nodes to N_{np}
9: Class all outputs of N_{np} as non-productive canopy

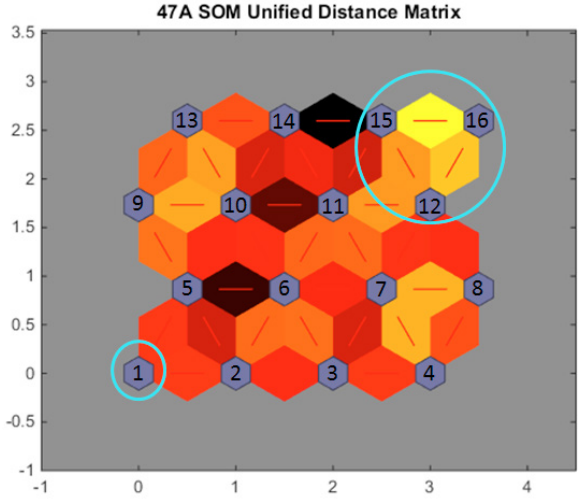

Fig. 4. SOM unified distance matrix output for block 47A. Lighter colors represent shorter distances between nodes. Nodes 1 and 16 are the dominant nodes representing the non-productive class. Nodes 12 and 15 are closely related to Node 16 and are therefore also considered as part of the non-productive class (highlighted in cyan).

$$p = \begin{cases} \text{green} & \text{if } H \in [80^o, 120^o] \wedge S > 20\% \wedge V > 50\% \\ \text{not} & \text{otherwise} \end{cases} \tag{8}$$

An alternate method to partition green pixels was also considered by mapping pixels in each frame to the L*a*b color space, with the green partition condition given by (9).

$$p = \begin{cases} \text{green} & \text{if } a < \frac{120}{255} \wedge b > \frac{136}{255} \\ \text{not} & \text{otherwise} \end{cases} \tag{9}$$

4. RESULTS AND DISCUSSION

The experimental procedures in this paper were evaluated in their performance using the following metrics of average precision and average recall.

$$\text{average precision} = \frac{\sum_{n=1}^{n_{rows}} \frac{TP_n}{TP_n + FP_n}}{n_{rows}} \tag{10}$$

$$\text{average recall} = \frac{\sum_{n=1}^{n_{rows}} \frac{TP_n}{TP_n + FN_n}}{n_{rows}} \tag{11}$$

Table 1. Variation in missing vine classification from aerial maps

Block	Vine spacing threshold	Avg Precision (row)	Avg Recall (row)
40A	0.5	0.9643	0.8492
40A	0.25	0.9167	0.8988
40A	0.2	0.8929	0.8988
40A	0.15	0.8651	0.8988
47A	0.5	0.7736	0.6150
47A	0.25	0.3574	0.9279

The results shown in Table 1 provides information about the variation in ability to identify the number of missing vines in each vineyard row. By varying the threshold indicating the amount of vine spacing required before being classed as 'missing vine', the average precision and recall varies.

It can be seen that using a 50% vine spacing threshold performs best due to the higher precision and recall when compared to smaller vine spacing thresholds. This is especially the case for block 47A which contains many thin or unhealthy vines. Bare-wire between such vines are not actually missing vines. On the other hand, block 40A has similar performance when using 50% and 25% vine spacing threshold. This can be attributed to much healthier vines which grow into the space of missing vines. Thus, we conclude that the health of vines in the vicinity of a bare-wire region needs to be considered when selecting a vine spacing threshold.

Table 2. Variation in feature settings and effect on average precision and recall of non-productive canopy calculated on a row basis from aerial data

Block	Feature Omitted	Precision (row)	Recall (row)
40A	none	0.8668	0.6089
40A	NDVI	0.716	0.7887
40A	NDVI-Blue	0.8411	0.6483
40A	NDVI-Green	0.8832	0.5158
40A	Green-Red	0.8728	0.5457
40A	Green-Blue	0.8886	0.5498
40A	PCD	0.9487	0.3495
47A	none	0.8649	0.7087
47A	NDVI	0.882	0.7122
47A	NDVI-Blue	0.8759	0.7298
47A	NDVI-Green	0.8623	0.7264
47A	Green-Red	0.9626	0.5261
47A	Green-Blue	0.8933	0.631
47A	PCD	0.8675	0.7070

The results for blocks 40A and 47A are shown in Table 2, where the contribution of each feature used in the classification of non-productive canopy to the average precision and recall for each vine row can be seen. Although our results are based on using all features, it may be possible to reduce the feature space size by omitting features with low impact, so as to reduce training time and increase accuracy and precision.

The omission of NDVI-Blue from the feature vector increases the overall recall while only slightly reducing preci-

sion, for block 40A. The results shown in Table 3 reflect the percentage of non-productive canopy in each block based on the omission of NDVI-blue. The results for B4 and B12 are only estimates as there is no ground truth available to compare the results to.

Table 3. Total percentage of non-productive canopy on a block level and average precision and recall values derived from the feature set (with omission of NDVI-blue)

Block	Manual Counts	Estimated	Precision	Recall
40A	3.77%	3.56%	0.8411	0.6483
47A	18.48%	15.30%	0.8759	0.7298
B4	-	0.76%	-	-
B12	-	2.23%	-	-

It is evident that the percentage of non-productive canopy on a block basis is similar to that estimated from manual counting; allowing this parameter to be used in models for decision support or yield estimation. Although the precision and recall appears to indicate that localization estimates of all non-productive canopy, as shown in Fig. 6 may not be accurate with the approach presented in this paper, it is worth noting that the ground-truth was collected at a very early E-L stage.

For block 40A, manual counts were collected at E-L stage 4. At this particular E-L stage, the extent of canopy needs to be estimated as canopy has not fully developed. Thus, this can lead to a high variation between the manual counts and the aerial imagery processed results. For block 47A, a higher recall was achieved which could be due to manual counts being obtained at a later E-L stage (7). As canopy had slightly developed further, a smaller amount of bias in estimating canopy extent may be obtained.

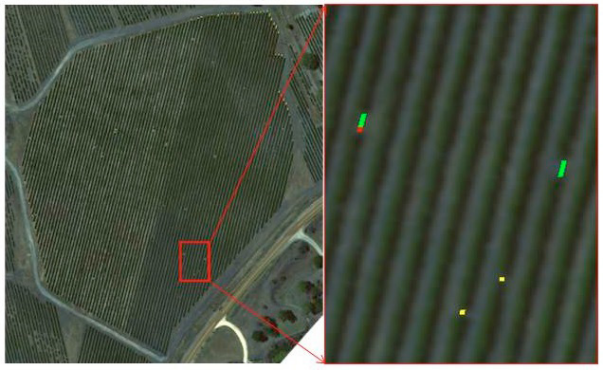

Fig. 5. Bare-wire in block 40A. Patch of block highlighted in red. Section of bare-wire in block 40A. Bare-wire regions highlighted in blue and missing vines are highlighted in red or green.

Fig. 7 and 8 show the proximal classification of green pixels in video frames. Missing rows are a result of substandard video recordings being discarded due to inconsistent lighting throughout recording of that row. Table 4 shows the precision, recall and estimated non-productive percentage of each block. Here, the non-productive percentage is a fairly accurate measure of the manual non-productive measurements for 40A, whereas the percentages overestimate non-productive proportions for 47A. Identical green pixel thresholds were used for both blocks and it may be the case that the correct range of *green* is slightly different between the two cultivars. Although the estimated non-productive percentages are favorable for 40A, the spatial maps have little correlation. Given the proximal data was collected during the shoots stage of growth and the aerial photography taken during veraison, further investigation is required to determine if aerial and proximal methods are more appropriate at a given phenological stage.

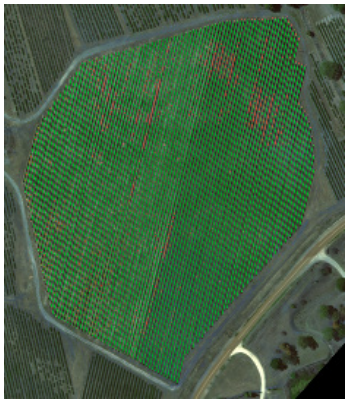

Fig. 6. Non-productive canopy in block 40A produced from aerial data. Non-productive vines highlighted in red.

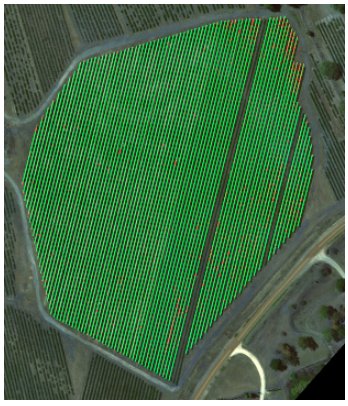

Fig. 7. Spatial map generated from equation (8) thresholding.

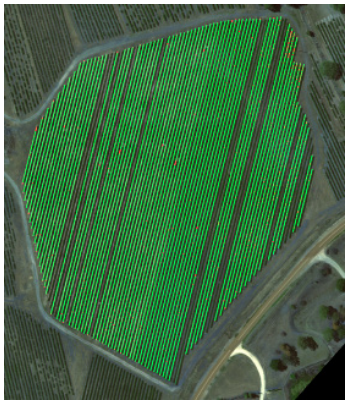

Fig. 8. Spatial map generated from equation (9) thresholding.

Table 4. Non-productive canopy results from proximal data showing average precision and recall on a row basis as well as estimated non-productive canopy estimates on a block basis

Block	Non-productive (%)	Precision (row)	Recall (row)
40A-LAB	1.26	0.9735	0.3313
40A-HSV	3.75	0.8388	0.592
47A-LAB	25.36	0.8409	0.4256
47A-HSV	37.02	0.7708	0.4761

5. CONCLUSION

This paper presented a method for localizing and classifying missing vines and non-productive canopy at two different phenological stages. Results in missing vine classification from aerial imagery indicate that using vine spacing as a thresholding parameter for identifying the number of missing vines in a block does not perform well across different blocks as it is dependent on the health of surrounding vines. Thus, future work will take into account such spatial parameters.

In non-productive canopy measurements, it was found that localization of non-productive canopy is difficult in both aerial and proximal approaches. However, promising results were obtained for estimating the percentage of non-productive canopy on a block basis from aerial imagery obtained around veraison and from proximal video data for Chardonnay at the early E-L stage of shoots. This can therefore be utilised for as a parameter for decision support or yield estimation models without the requirement of labor intensive manual measurement.

The advantage of the proximal approach is early season sensing of non-productive canopy which is unlikely to be feasible with aerial imagery from commercial systems due to limitations in spatial resolution of imagery when dealing with shoots. Thus, these two methods can be applied in a complementary manner for estimating non-productive canopy at an early stage and verifying at a later stage.

Overall, this work seeks to assist in improving management practices by providing additional information about spatial variation of non-productive canopy in an automated manner.

ACKNOWLEDGEMENTS

The authors would like to thank Angus Davidson for assistance with ground truth data collection and insight on non-productive canopy. Many thanks to Jarrett's of Orange and Treasury Wine Estates for provision of facilities and data critical for the conducting of this research and Wine Australia for financial support through project DPI 1401, "Improved Yield Prediction for the Australian Wine Industry." The first author is also partly supported by Wine Australia grant AGW Ph1513, "Autonomous In-Vivo Determination of Maturity Parameters."

REFERENCES

Calderón, R., Navas-Cortés, J.a., Lucena, C., and Zarco-Tejada, P.J. (2013). High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sensing of Environment*, 139, 231–245.

Comba, L., Gay, P., Primicerio, J., and Ricauda, D. (2015). Vineyard detection from unmanned aerial systems images. *Computers and Electronics in Agriculture*, 114, 78–87.

Coombe, B. (1995). Growth stages of the grapevine: adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research*, 1(2), 104–110.

Hall, A., Lamb, D.W., Holzapfel, B.P., and Louis, J.P. (2011). Within-season temporal variation in correlations between vineyard canopy and winegrape composition and yield. *Precision Agriculture*, 12(1), 103–117.

Hall, A., Louis, J., and Lamb, D. (2003). Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images. *Computers & Geosciences*, 29(7), 813–822.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464–1480.

Liu, S., Tang, J., Cossell, S., and Whitty, M. (2015). Detection of shoots in vineyards by unsupervised learning with over the row computer vision system. In *Australasian Conference on Robotics and Automation*.

Llorens, J., Gil, E., Llop, J., and Queraltó, M. (2011). Georeferenced LiDAR 3D vine plantation map generation. *Sensors*, 11(6), 6237–6256.

Mathews, A.J. and Jensen, J.L.R. (2013). Visualizing and quantifying vineyard canopy LAI using an unmanned aerial vehicle (UAV) collected high density structure from motion point cloud. *Remote Sensing*, 5(5), 2164–2183.

Nolan, A.P., Park, S., Fuentes, S., Ryu, D., and Chung, H. (2015). Automated detection and segmentation of vine rows using high resolution UAS imagery in a commercial vineyard. In *MODSIM2015, 21st International Congress on Modelling and Simulation*, 1406–1412.

Nuske, S., Gupta, K., Narasimhan, S., and Singh, S. (2014). Modeling and calibrating visual yield estimates in vineyards. *Springer Tracts in Advanced Robotics*, 92, 343–356.

Sauvola, J. and Pietikäinen, M. (2000). Adaptive document image binarization. *Pattern recognition*, 33(2), 225–236.

Serrano, L., González-flor, C., and Gorchs, G. (2012). Remote Sensing of Environment Assessment of grape yield and composition using the re fl ectance based Water Index in Mediterranean rainfed vineyards. *Remote Sensing of Environment*, 118, 249–258.

Turner, D., Lucieer, a., and Watson, C. (2011). Development of an Unmanned Aerial Vehicle (UAV) for hyper resolution vineyard mapping based on visible, multispectral, and thermal imagery. *Proceedings of 34th International Symposium on Remote Sensing of Environment*.