Australian Government

Wine Australia

CSIRO

Australian Government

Department of Agriculture, Water and the Environment

Rural R&D for Profit Program

New non-destructive technologies for simultaneous yield, crop condition and quality estimation

Final Report - CSA 1602

CSIRO Agriculture & Food

October 2016 – March 2020

Everard J Edwards



New non-destructive technologies for simultaneous yield, crop condition and quality estimation

CSA1602

October 2016 to March 2020

Everard J Edwards, Mark Thomas, Stephen Gensemer, Ryan Lagerstrom, Rizwan Khokher, Qiyu Liao, Changming Sun, Dadong Wang, Chad Hargrave, Jonathon Ralston

CSIRO
Locked Bag 2
Glen Osmond
SA 5064

COPYRIGHT

© 2020 CSIRO. To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

DISCLAIMER

CSIRO Agriculture & Food advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice.

To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

Contents

Plain English summary **1**
Abbreviations and glossary **3**
1 Project rationale and objectives **4**
2 Method and project locations **6**
2.1 Methods 6
2.2 Locations 15
3 Project outcomes **17**
3.1 Project level achievements 17
3.2 Contribution to program objectives 46
4 Collaboration **48**
5 Extension and adoption activities **49**
6 Lessons learnt **50**
7 Appendix - additional project information **51**
7.1 Project, media and communications material and intellectual property 51
7.2 Equipment and assets 51
7.3 Thermal Infrared Imaging – Initial Evaluations 52
7.4 References 53
7.5 Staff 54
7.6 Monitoring and evaluation 55
7.7 Budget 55

Plain English summary

Accurate yield estimation is a key part of horticultural production, with supply chains and fruit value significantly impacted when production volume doesn't match expectations. The wine industry is no different in this respect, with everything from winery intake planning to sales reliant on yield estimation to some extent. Current yield estimates rely significantly on the experience of the vineyard manager, are labour intensive and typically have an error rate of up to 30%. Improvements in yield prediction and estimation offer significant value to the industry.

More than any other perennial crop, the commercial potential of winegrapes depends on composition, with the same variety fetching a ten-fold difference in price depending on its quality. Objective assessment of fruit quality is, therefore, paramount. If this can be done in the field prior to harvest and transport, winery planning can be optimised, a number of potential points of friction between grower and winery can be avoided and additional opportunities would be available to the grower to produce fruit to a desired specification. Similarly, accurate and objective pre-harvest assessment of crop condition would minimise winery rejection and provide the best opportunity for a profitable harvest.

With the recent revolution in sensors and data analytics, together with a pervasive online computing infrastructure to support these, there is an opportunity to provide the grape-grower with cost effective tools to predict inter-seasonal variation yield from early in the season, to monitor fruit ripening and quality during maturation and to estimate yield and fruit condition prior to harvest. Furthermore, these tools can be based on whole vineyard assessments using non-destructive measurements, rather than few, small, labour intensive, destructive samples.

This project was established to test and develop technologies to provide non-destructive, non-contact yield prediction, yield estimates, monitoring of fruit composition and in-field estimates of fruit condition. Colour digital video, stereo imaging, and low-power radar were all used to collect in-field data for yield prediction and estimation. All these devices were mounted on commercial vineyard vehicles and used to image commercial vineyards in collaboration with a number of wine companies (primarily Accolade Wine, but also Katnook Estate & Rymill Wine). Bespoke data-processing pipelines were developed in all cases to produce grower relevant data from the sensor outputs.

A lab-based hyperspectral imaging system was used to develop detailed calibrations of fruit composition (quality) against several thousand individual lab analysed samples. This same system was trialled as a field device, but in addition, a cheaper and more robust field system was developed using dual point-spectrometers and the same data analytics.

The project has clearly demonstrated the viability of using colour digital video and machine learning to count inflorescences at scale in the vineyard and thereby provide a yield prediction shortly after budburst. The same method works reliably for bunch counting on green or red grapes, but needs further optimisation for vineyards with heavier canopies. Low-power radar was able to detect fruit irrespective of canopy, but its ability to reliably quantify fruit mass still needs to be demonstrated.

Hyperspectral data analytics were able to accurately quantify key fruit maturity and quality parameters such as sugar content and total organic acids, across a broad range of winegrape

varieties and fruit development stages, but estimates of berry nitrogen parameters were only effective on a per genotype basis. Although the calibrations were lab based, their use with field imaging was demonstrated and field devices to collect the relevant data developed and trialled.

It is recommended that each of these technology areas be assessed, with further industry input, for commercialisation potential, particularly considering the dollar value of the data to viticulturist decision making. For example, if the industry values a key parameter as being cost effective at $50 per hectare, then the potential to produce that data at that cost needs to be demonstrated.

For those systems at a lower technology readiness where this assessment it not yet practical, it is recommended to continue research where there is ongoing industry support.

This project was supported by Wine Australia, through funding from the Australian Government Department of Agriculture, Water and the Environment as part of its Rural R&D for Profit program, and CSIRO.

Abbreviations and glossary

2.5D – 2-dimensional imaging with depth

FoPEN – foliage penetrating radar

GPU – graphical processing unit

HSI – hyperspectral imaging

LASSO – least absolute shrinkage and selection operator

NIR – near infra-red radiation

PLS – partial least squares regression

RF – random forests

RGB images – red-green-blue images (standard digital imaging)

RMSE – root mean square error

TIR – thermal infra-red

TSS – total soluble solids

1 Project rationale and objectives

Yield prediction and estimation have long been a key target for viticultural research and high priorities for the wine industry. At present, yield estimation is a subjective assessment with very small sampling sizes compared to vineyard size. This commonly results in errors up to 30%, resulting in an under or over supply to the winery. Knowing the yield accurately ahead of harvest allows a winery to better plan intake and adjust for over or under supply of a variety. Currently, yield estimation costs the industry around $16/ha and the high error rates enlarge this cost. A reduction in the yield estimate error from 33% to 20% would amount to a saving of $85 million per year (Clingeleffer et al., 2001). Further reductions would magnify the savings.

Assessment of fruit condition and quality is another ongoing industry issue. Not only in terms of the cost of monitoring, fruit streaming and scheduling, but also in disputes between growers and wineries on price paid and fruit rejection. Meeting winery specifications can be a contentious issue and was recently the subject of an Australian Competition and Consumer Commission investigation. This is exacerbated when subjective assessments of fruit condition and quality determine grade and price. For example, as little as 5% of berries infected with fungus or 10% damaged fruit can lead to rejection of a crop. Achieving an accurate assessment of fruit condition and quality would also significantly benefit operations for both growers and wineries.

These industry problems are due to the limited sampling/observations that are practical within a vineyard and the lack of objective data for many assessments. The labour cost involved in monitoring and undertaking these assessments is also an industry concern.

The ideal solution to this problem is accurate objective measurements of yield, fruit condition and composition (quality) that are done in the vineyard at appropriate time points during maturation. This would also permit the identification and separation of berry damage pre-harvest and after harvest and how winery-imposed penalties, if any, might be apportioned between the grower and harvester.

Ideally these objective measurements would not be spot sampling but cover the whole vineyard. Also essential are methods for easy, fast data collection in the vineyard, fast processing and analysis of the data and delivery of this information to the grower and winery in an acceptable format and timeframe.

Understandably, yield prediction for winegrapes has been an active research area in Australia and overseas for a long time. The biology of the crop and management methods to manipulate yield have been extensively, albeit not exhaustively, researched, however, accurate yield prediction is a problem still not solved. Perhaps the most encouraging development was a paper by Nuske et al., (2014) using 2D digital imaging with artificial lighting in US vineyards at night, reporting an 89-97% accuracy just prior to harvest.

Our project aimed to improve on this approach by evaluating new non-contact, non-destructive technologies that can be deployed at scale providing accurate, fast measurements of yield, crop condition and quality information at relevant time points and in acceptable timeframes that assist in decision-making for both the grower and winery.

These technologies included 3D imaging for yield prediction during the day; foliage penetrating radar (FoPEN), already used to estimate yield in sugar beet while the beet is still in the ground (Konstantinovic, et al., 2008) and hyperspectral imaging (HSI), which has been finding increasing application in agriculture (e.g. disease detection (Mahlein, 2012), fruit sorting (Gruna, 2010)).

Further, the project aimed to take a high-performance computing (HPC) based approach to the data analysis, with data storage and processing to be in the "cloud". This was to maximize accessibility for end users and to ensure that the adoption pathway will be a viable solution for industry in the future. Industry would use mobile devices (smart phones, tablets, laptop computers) or computers to access the information via a cloud server service for decision-making purposes.

The project objectives remained largely unchanged for the duration of the project and were:

- To evaluate new vineyard methods and sensors for the collection of
    - yield data from vineyards using novel imaging hardware (FoPEN and 3D imaging) at different time points prior to harvest and,
    - bunch condition and quality data (digital imaging and HSI).
- To develop and evaluate new data analysis pipelines and algorithms for:
    - rapid feature extraction to identify bunches and inflorescences to determine their number and estimate volume for yield estimation and,
    - identification of spectral signatures and features that provide a quantitative estimate of bunch condition (fruit damage (shrivel) & disease) and berry quality (Brix, colour, acid).

The intention of the project was to develop tools that could be applied to the full range of Australian viticulture, from high value iconic cool-climate vineyards to high yielding vineyards in the warm irrigated regions. This objective and non-destructive approach intended to assist the industry by providing the following benefits: forward planning for winery intake, reduced disagreement between growers and wineries on crop price, reduced need for subjective assessment of fruit and improved understanding of differences between high value vineyards and lower value vineyards for implementing change to lift value

2 Method and project locations

The major objectives of the practical work were i) to develop tools to non-destructively predict yield at scale at different time-points prior to harvest, ii) develop hyperspectral calibrations for key factors determining fruit grade, such as sugar content, colour (in red grapes), acidity, etc. and test under field conditions; and iii) assess the potential for hyperspectral imaging to be used for assessment of fruit condition, e.g. proportion of berry shrivel or disease, in the field. These objectives utilised a combination of cost-effective consumer products, off-the-shelf scientific grade equipment and new hardware developed by CSIRO.

2.1 Methods

The key technologies used during the project were radar, RGB digital imaging and hyperspectral imaging. RADAR (Radio Detection and Ranging) is a technology invented in the 1940's for detecting the presence and movement of large metallic (electromagnetically conducting) objects such as aeroplanes, which uses a single antenna system to transmit pulses of radio and receive the echo from an object. More recently, low power radar-on-a-chip devices have become available that work over the cm to m range. Even low-power, high frequency radar has the ability to pass through some objects and the contrast in conducting volume (water) of a grape bunch to a leaf offers the opportunity to use low-power radar in the vineyard, post-veraison, to detect and estimate fruit mass, even where the canopy obscures the bunches.

Hyperspectral imaging refers to the collection of a spectrum (many wavebands) for each pixel of an image. Hyperspectral sensors have been incorporated into satellites and used in aerial imaging for over 20 years. In agriculture, such systems have been investigated for disease detection, e.g. in sugarcane (Apan et al. 2004), crop growth parameters (Patel et al. 2001, Backhaus et al. 2011) and even identification of grapevine cultivars (Lacar et al. 2001). More recently, improvements in the sensitivity of sensor systems have led to use across widespread areas of research and industry alike (Mazzetto et al. 2010).

Together with the RGB digital imaging, familiar to all, these technologies were used as follows.

2.1.1 Yield prediction and estimation

Vineyard yield is a function of the number of vines, the number of grape bunches per vine and weight of those bunches. Bunch weight is, in turn, a function of the number of berries per bunch and the berry weight. As the number of vines, ignoring any vine losses, in a vineyard is fixed, the year to year variation in yield is a function of the change in bunch number, berry number and berry weight. Bunch number is the largest factor, typically accounting for 60% of inter-seasonal yield variation (Clingeleffer 1997), with berry number per bunch and berry weight both accounting for around 20% of yield variation. Bunch number is initially set by inflorescence number, which is the reason behind undertaking bud dissection to determine fruitfulness of vines. This project aimed to provide yield prediction as early as possible in the season. Whilst using Near Infra-Red (NIR) spectral information has potential for assessing fruitfulness of buds, it is unlikely to be developed into an on-the-go non-contact sensor and was being examined, using an active sensor, in a separate RRD4P project. Consequently, the earliest potential yield prediction that could be examined using proximal sensing was shortly after budburst when the inflorescences were clear of the developing shoot, at the E-L 12 growth stage (Coombe 1995).

The project therefore aimed to be able to estimate bunch/inflorescence number at various phenological stages from E-L 12 onwards and use this to predict yield based on *a priori* knowledge of typical yield from the block utilised. In addition, we examined methods to estimate yield or the other yield components directly, by determining fruit mass.

RGB video and machine learning: a major component of this project was the development of techniques to use RGB video from low cost cameras for yield prediction and estimation. Beginning early in the 2017/18 season and continuing through to the 2019/20 season, Go-Pro consumer action cameras (with in-built GPS) were mounted at various combinations and angles (Figure 1 provides the most comprehensive set-up) on the Kubota RTV vehicle. Videos were primarily taken at the Accolade Mclaren Vale vineyard and typically of one side of a row at a time, although continuous videos of multiple rows were taken in the 2019/20 season to trial photogrammetry tools. Additional data were collected from the Adelaide Hills vineyard at Woodside in 2018/19 and 2019/20. Videos were taken in a variety of weather conditions, but not when precipitation was occurring.



Figure 1 – Go-Pro array for vineyard row imaging mounted onto vineyard vehicle. Number of cameras mounted and used for a given run varied.

The data collection was targeted at key phenological stages: post-budburst (E-L 12-14), veraison (E-L 35), mid-maturation (E-L 36) and harvest (E-L 38), with the aim of capturing inflorescence number, bunch number and, potentially, bunch size and berry number.

Videos were split into individual frames and a sub-set for manual labelling selected to represent the row (i.e. approximately evenly spaced). Labelling was undertaken using the QuPath tool (Bankhead et al., 2017) and consisted of drawing a bounding box (rectangle) that encompassed each object of interest (inflorescence or bunch) within an image (e.g. Figure 2). Labelled images were used for a deep learning approach, which used a ResNet backbone with a RetinaNet detector to learn the location of the objects of interest, outputting a bounding box. Once trained, the detector could be used to analyse and locate the objects in any number of video frames.



Figure 2 – cropped section of video frame with inflorescences marked by hand using a bounding box to encompass the feature.

However, as the video frames had a high degree of overlap a method to ensure all bunches/inflorescences were not only counted, but only counted once, was required. To achieve this a tracking algorithm was developed that located each object and tracked it through adjacent video frames. This involved a number of steps, with all possible paths between detections in adjacent frames being determined, weighted, then selected automatically using computer-based image analysis techniques. In some circumstances, objects would be detected in some frames, not appear in several frames, then reappear later on, either due to glare from the sun or occlusion from leaves. To counter this, an additional machine learning step was used to link broken trajectories of the same object. Figure 3 provides a starting image and subsequent trajectories of detected objects. The trajectories are not straight lines due to movement of the field vehicle and the curvature induced in the images by the camera lens.

Performance of the detector was assessed against the digital ground-truth (the manually labelled images) and the performance of the combined detection and tracking algorithm against hand counted ground-truth values from the vineyard.

A second method to eliminate multiple counting of objects was also investigated. This was to produce a single large mosaic image for each row from the video, whereby each object would

only appear once. To achieve this, the rough overlap for neighbouring images in an image sequence was estimated. Then mosaic sub-images were then selected and matched with neighbouring overlapping sub-images (Sun 2002). The information in the matched image was used to merge the overlapping images. The mosaicking process was a modified version from the approach described in (Sun et al., 2006). For any particular mosaic, the building blocks were from vertical slices of the input images which represented a narrow view angle and not the entire field of view of the camera (approx. 88 degrees). Consequently, different mosaics could be generated, which had different view angles.



Figure 3 – single video frame with trajectories if selected inflorescences through subsequent frames marked.

Stereo imaging: this was undertaken using a pair of Grasshopper3 GigE machine vision cameras (GS3-PGE-91S6C-C, Flir, formally Point Grey, Edmund Optics, Singapore) with the aim of obtaining 2.5D (2D RGB image+depth) data. The advantages of this were two-fold. Firstly, knowing the distance to a bunch in an image allowed it to be scaled, thereby having an estimate of bunch size. Secondly, high resolution depth data offered the potential to produce a 3D model of bunch shape and thereby estimate bunch volume.

The cameras were controlled using a laptop running Ubuntu Linux 16.04 together with an external trigger to ensure pairs of images were captured simultaneously. Each resulting image was 3376×2704 pixels. The camera system was mounted on a frame angled up towards the bunch zone at about 45 degrees and this, in turn, was mounted on a vineyard vehicle. The vehicle speed was approximately 6-8 km/h and images were collected at 9 Hz. The hardware setup is given in Figure 4. The pipes, angled backwards, were used to keep the vine canes away from the cameras, particularly at later growth stages.

Stereo imaging was undertaken pre-harvest in the 2016/17 and 2017/18 seasons.



Figure 4 – stereo machine vision camera set-up, mounted to the vineyard vehicle with protective frame (left) and close-up with supplementary light-sources to limit constrast caused by direct sunlight in parts of the images.

Radar: fruit are rarely fully exposed to the sun in Australian vineyards and 'sprawl' canopies are common to protect fruit from direct sunlight during the hotter parts of the season. The limited ability of visible radiation (or NIR) to penetrate the canopy, means that ways must be found to provide a direct view of the fruit. However, other wavelengths do have the power to penetrate solid materials to an extent that might allow fruit to be imaged. High-energy shorter wavelengths, such as x-rays, can do this, but have safety concerns and a lack of suitable field equipment. Portable x-ray fluorescence instruments are available, but are not suited to the task and have potential safety issues. On the other hand, longer wavelengths also have the ability to penetrate through material such as leaves and have few safety risks. Low-power radar in particular offers the potential to see through canopies and quantify the fruit mass within them.

An ultra-wide band device consisting of an array transmitters and receivers, together with RGB imaging, was built and mounted on the field vehicle and tested at the Accolade vineyard in Mclaren Vale during the 2016/17 season (Figure 5 top), with leaves and fruit being stripped in sequence from a panel and multiple scans being made and demonstrated that fruit provided stronger returns than leaves.

Following this computational electromagnetic modelling was undertaken for bandwidth of 0-8 GHz. This progressively built a section of vine with berries, leaves and canes to simulate the radar returns for further hardware optimisation. A software tool to visualise and explore the radar data together with the RGB images was also developed and improved as the project continued (Figure 5 middle). A test 'rack' was also built to allow various combinations of canopy and fruit at known distances to scanned by the radar devices for testing (Figure 5 bottom). Various iterations of the hardware were tested using the rack and in the field over the subsequent seasons.



Figure 5 – low-power radar system in use in the vineyard (top); radar and RGB camera output in software tool (middle); and lab-based test rig with design (bottom).

2.1.2 Non-destructive measurement of fruit composition

Analysis of fruit spectral reflectance was considered the only realistic non-destructive, non-contact method to determine fruit quality (composition). Specifically, hyperspectral imaging (HSI) in the NIR was used with various statistical and machine learning tools to build calibrations against key components of the fruit. The approach used was to demonstrate the potential of HSI to measure fruit composition under optimal, lab-based, conditions and then to develop an on-the-go system capable of measuring those same components in the field.

The laboratory-based work used the same system described in the final report for CSA1601, a Headwall Micro eVNIR line-scan camera with a Headwall halogen light source (Portable Analytical Solutions, NSW, Australia) and Zaber linear stage (Edmund Optics, Singapore) to move the sample past the camera. The camera was capable of imaging 150 wavebands over a spectral range of 550-1650 nm and had 300 pixels per line. A tray to hold sixty separated berries and a white reference target was designed and 3D printed to mount on the linear stage. Due to the late arrival of the instrument it was not possible to collect berries during the 2016/17 growing season, so the initial testing was done on berries from four glasshouse grown microvine varieties. Fruit from both red and white varieties at different phenological ages, from pre-veraison to post-harvest, were scanned using the Headwall HIS system, following which the individual berries were crushed and the composition measured using an Oenofoss (Foss, Denmark). In addition, a sub-set of red berries were analysed for anthocyanin and phenolic content using a modified version of the assay by Iland et al. (2000).

A basic pipeline to automatically obtain the spectra per berry was developed. The spectra were corrected to reflectance using a dark and white reference per tray, berries were separated from the background by subtracting the intensity of the 670 nm waveband from the 760 nm band, connected components were identified to determine individual berries, following which specular reflection of the lamp from the curved berry surface removed and, finally, the pedicel identified and removed (Figure 6). Typically, the median spectra per berry was then determined and used to build a calibration against Oenofoss or other wet lab results using a variety of statistical and machine learning tools.

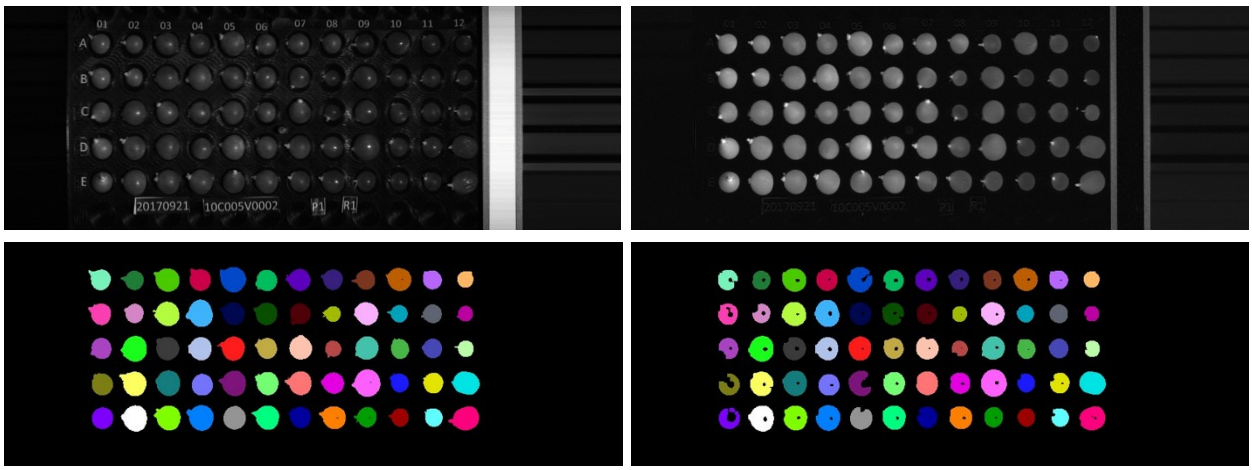


Figure 6 – processing pipeline for HSI imaging of berries: raw scan at 984 nm (top left), difference between 761 and 670 nm after reference correction (top right), determination of individual berries (bottom left) and pixels per berry selected after removal of pedicel and specular reflection (bottom right).

Following the initial tests with glasshouse berries in 2017 (n = 270 for red berries & 740 for white berries) further data sets were collected during maturation in the 2017/18 growing season, with weekly sampling of Shiraz berries from the University of Adelaide Waite campus vineyard from shortly before veraison to harvest (12 dates); during maturation in the 2018/19 growing season, with weekly sampling of Sauvignon Blanc berries from the Waite Campus vineyard (nine dates); and during the 2019/20 growing season with sampling at three time-points of Chardonnay, Semillon, Taminga, Muscat Hamburg, and Carignan berries, again from the Waite Campus vineyard. In total, 7,280 berries were scanned and individually measured for juice composition.

The next step in this work was to move the data collection to on-the-go measurements in the field. Due to the differences between solar radiation, with a number of atmospheric absorption bands reducing the effective wavelengths available, and the broad-spectrum lab illumination, either a light source was required with the HSI or the calibrations needed to be adapted to work in sunlight. Further, rather than the sample being moved past the linescanner at a known and controlled speed, with no other movement, the instrument now needed to be moved past the target on a vehicle that would have some degree of movement in all three dimensions, relative to the target. Finally, the target itself could be moving due to wind.

The first attempt to do HSI under sunlit conditions in the field was undertaken in the 2016/17 growing season using a different instrument, a low-cost Ximea MQ022HG-IM-LS100-NIR camera (Ximea GmbH, Münster, Germany). Although this camera was a linescan sensor, it used simultaneous data collection over a 2D chip, rather than a push-broom system. Each waveband (100 bands between 600 and 1000 nm) was collected simultaneously, but had a different view, with movement required to build up the spectra rather than the image.

Due to movement between frames, manual tracking was required to assemble a full data-cube (2D image with full spectra per pixel). Alternative approaches, such as blurring adjacent rows in non-tracked images and selecting areas of interest allowed spectra to be recreated for objects without the manual tracking, but lost 2D shape information. Ultimately, this camera, which was a fraction of the price of the push-broom style Headwall linescan camera and therefore more viable for commercial uptake, was deemed unsuitable for field use.

For field use, the Headwall camera was mounted together with an RGB video camera and a high-power xenon flash (Figure 7) and mounted to the Kubota RTV field vehicle for simulated commercial field use. Testing was carried out at Accolade Mclaren Vale on a small number of vine panels, with bunches harvested after imaging and analysed for TSS.

An HSI instrument such as the Headwall is unlikely to become viable commercial for commercial use in the near future due to cost. However, the results from the tissue classification undertaken with the Headwall in the CSA1601 sister project demonstrated, together with the field tests described above, demonstrated that spectra could be identified as belonging to fruit, or other vine parts, and thus 2D imaging was not necessarily required to obtain field data from fruit. Consequently, a new, lower-cost, instrument was built, based on two point-spectrometers (C13053MA and C14486GA, both Hamamatsu, Shizuoka, Japan). In combination the two spectrometers had the same spectral range as the Headwall camera.

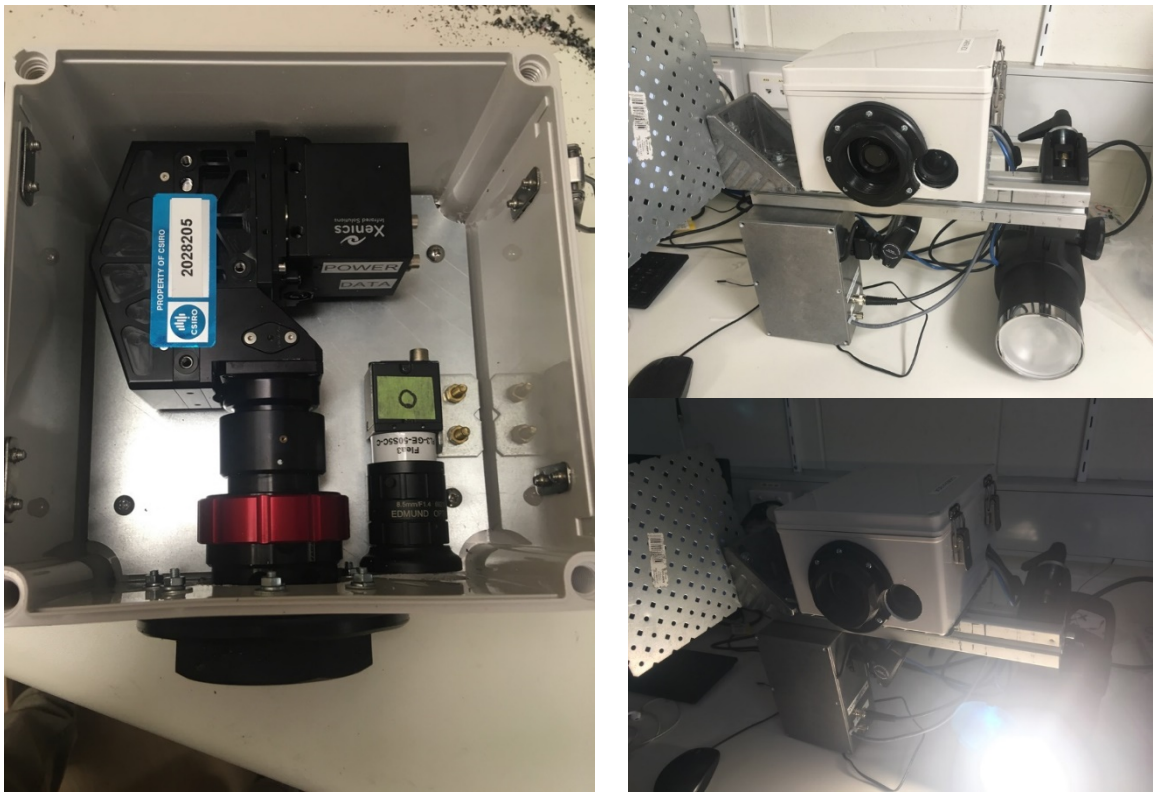


Figure 7 – Heawall eVNIR HSI camera mounted in field housing with paired RGB machine vision camera (left) and with high output xenon flash in use (right).

The new instrument was tested in combination with the Headwall camera using the lab set-up to enable direct comparisons, in the glasshouse, mounted to a linear stage with high-output halogen spotlights and in the field, mounted on the Kubota RTV vehicle, again using the high output halogen spotlights (Figure 8).



Figure 8 – twin point spectrometer based field instrument mounted on the vineyard vehicle together with halogen light source.

2.1.3 Non-destructive measurement of fruit condition

Two methods of in-field non-destructive assessment of fruit condition were considered, HSI and RGB imaging with machine learning, such as was used for inflorescence detection (see above). The latter was scheduled to be tested in a field site used by another Wine Australia funded project (CSA1701-2.2) during the 2019/20 season. Unfortunately, the vineyard suffered from fruit set problems during that season and the trial had to be abandoned. This will be re-examined as part of the CSA1701-3.1 project.

The advantage of HSI imaging for assessing fruit condition is the potential to estimate and separate multiple problems. For example, the potential of HSI to assess relative amounts of mildew infection in leaves was clearly demonstrated in the sister project CSA1601 and fruit condition and contamination with foreign matter at winery intake were examined using HSI as part of another RRD4P Wine Australia project (AWR1601).

To determine the potential of HSI to detect issues other than fungal disease and to assess the proportion of damage in whole bunches, fraction of berry shrivel was chosen a test parameter. Individual Shiraz berries with some degree of shrivel were cut from bunches and scanned and processed as described above. A Random Forest classifier was trained on these shrivelled berries and the non-shrivelled Shiraz berries scanned previously. The classifier was tested by scanning approximately 50 whole bunches that had been manually assessed and classified into 0-20, 20-40, 40-60, 60-80 or 80-100% shrivel categories, with two bunches per category per date.

2.2 Locations

Development of sensor equipment, laboratory and desktop analysis work was carried out at a number of CSIRO locations:

- CSIRO Pullenvale, QLD - radar hardware and data processing development.
- CSIRO Waite Campus, SA – consumer action cam system development, lab based hyperspectral and matching wet lab leaf analysis.
- CSIRO Marsfield, NSW – stereo camera development and data analysis, hyperspectral data analysis.

Testing of equipment, methods and tools was undertaken across a range of vineyards, intended to be representative of the majority of Australian viticulture. At the Accolade, Rymill and Katnook sites, multiple contrasting blocks were utilised. The majority of the work was carried out at Accolade Mclaren Vale and Woodside.

Commercial vineyards:

- Adelaide Hills – SA
    - Woodside – private vineyard
- Langhorne Creek – SA
    - Cleggett's – private vineyard
- Mclaren Vale – SA
    - Accolade
- Coonawarra – SA
    - Rymill Wine

- Katnook Estate

Research stations/farms:

- Adelaide – SA
  - University of Adelaide Waite Campus

3 Project outcomes

3.1 Project level achievements

This project has encompassed a significant amount of work by a large, geographically dispersed, team across a broad range of technologies and industry targets for digital viticulture. The reporting of project achievements below is necessarily a brief summary of that work and the learning associated with it. The project has achieved its objectives of developing and testing technologies for predicting and estimating yield, non-destructive determination of fruit composition (quality) and proximal assessment of fruit condition. The work presented here has demonstrated the viability of each of the approaches tested and we suggest that this is followed up by further assessment of the potential of these technologies for commercialisation in partnership with the wine industry.

3.1.1 Yield prediction and estimation

Two basic approaches were trialled to estimate the amount of fruit in the vineyard. Firstly, digital video imaging was used in an attempt to image as much of the fruit as possible and, secondly, foliage penetrating radar was used to try estimate the mass of the fruit. In both cases, new data analytics needed to be developed and a data processing pipeline produced.

Digital imaging had the advantage that it needed low cost equipment and the basic tools for the data processing were available, but relied on the fruit being visible from the outside of the canopy, albeit multiple camera angles could be used. Further, calibrations were required to convert the fruit visualised to mass. FoPEN had the potential to directly measure fruit mass through the canopy, but required new hardware, new data processing tools and was an unproven method for this purpose.

As described above, the yield component with the largest inter-annual variation is bunch number and, in the absence of bunch mass data, could be used with knowledge of historical yield components to make a yield prediction. Consequently, the major effort of the digital imaging work was to achieve an accurate estimate of bunch number.

Inflorescence counting: over 800 individual video frames, from three different vineyard blocks, taken under various sky conditions and view angles were hand labelled for the location of inflorescences. A subset of these have been publicly released under a Creative Commons license (https://doi.org/10.25919/5de4546aeacce). Examples of the variety of hand labelled images are provided in Figure 9.

The efficacy of the detector developed can be assessed using two metrics, the proportion of hand labelled objects detected (recall) and the proportion of detected objects that were in fact the object of interest (precision). A perfect model would have a precision and recall of 1.0. In practice, when training the detector a balance needs to be achieved between false positives, which reduce precision, and false negatives, which reduce recall. Typically, as recall approaches 1, whereby all labelled objects are detected, the number of false positives will increase.



Figure 9 – examples of hand labelled video frames demonstrating the range of conditions experienced during inflorescence imaging, including clear sky, sun in view causing glare, low light and difficult background objects such as trees or adjacent vine rows.



Figure 10 – video frames from figure 9 with bounding boxes determined through the deep learning based inflorescence detector.

Image labelling and optimisation of the detector is continuing, to improve the range of vineyards and image collection conditions for which it is effective, but at the time of writing recall was over 85% and precision over 80% (Table 1). The relationship between precision and recall can also be seen in the curves presented in Figure 11. The real-world result can be visualised by comparing Figure 10, labelled using the detector output, with the hand labelling in Figure 9 and a combined hand and machine labelled image is provide in Figure 12.

Table 1 – sample output of inflorescence detector training, demonstrating the trade off number of true detections and false positives, together with Precision – Recall results. Selected number of training epochs highlighted in bold.

No. of Epochs	**No. of Detections (GT = 2,747)**	**Recall**	**Precision**
6	TPs: 2,373, FPs: 1,319	0.864	0.792
10	TPs: 2,343, FPs: 753	0.853	0.804
11	**TPs: 2,338, FPs: 681**	**0.851**	**0.804**
14	TPs: 2,338, FPs: 734	0.851	0.802

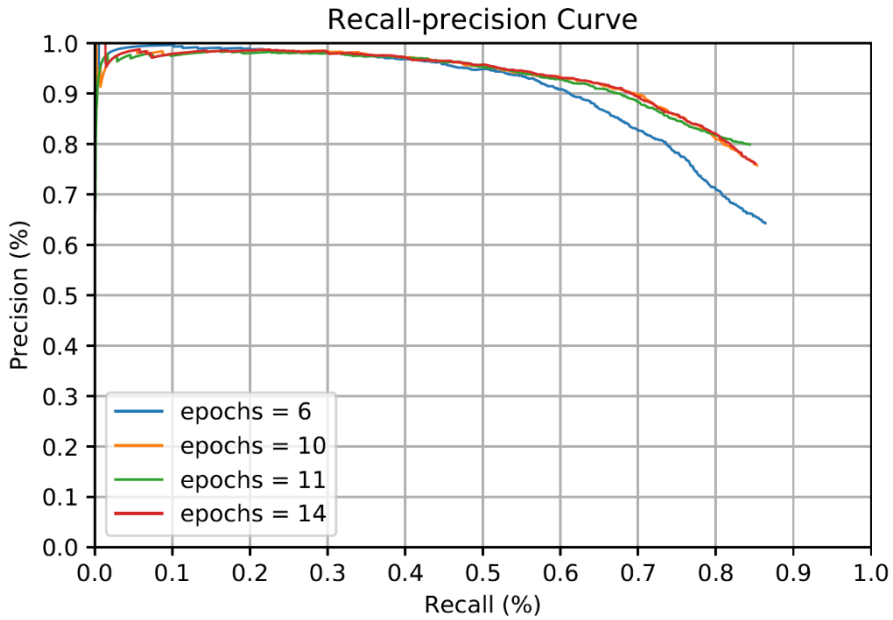


Figure 11 – visual representation of Precision-Recall curves for training epochs listed in Table 1.

Bunch counting: the same methodology was used to estimate bunch number later in the season. The work concentrated on two time-points, pre-veraison, whilst the berries were still green and, therefore, more difficult to differentiate from the canopy, and pre-harvest, when the fruit were fully coloured and the bunches at their largest. At both time-points the images from sprawl or loose VSP vineyards exhibited significant occlusion of bunches. Self-shading also sometimes made identifying bunches in the images difficult. In addition, the detected pre-harvest bunches

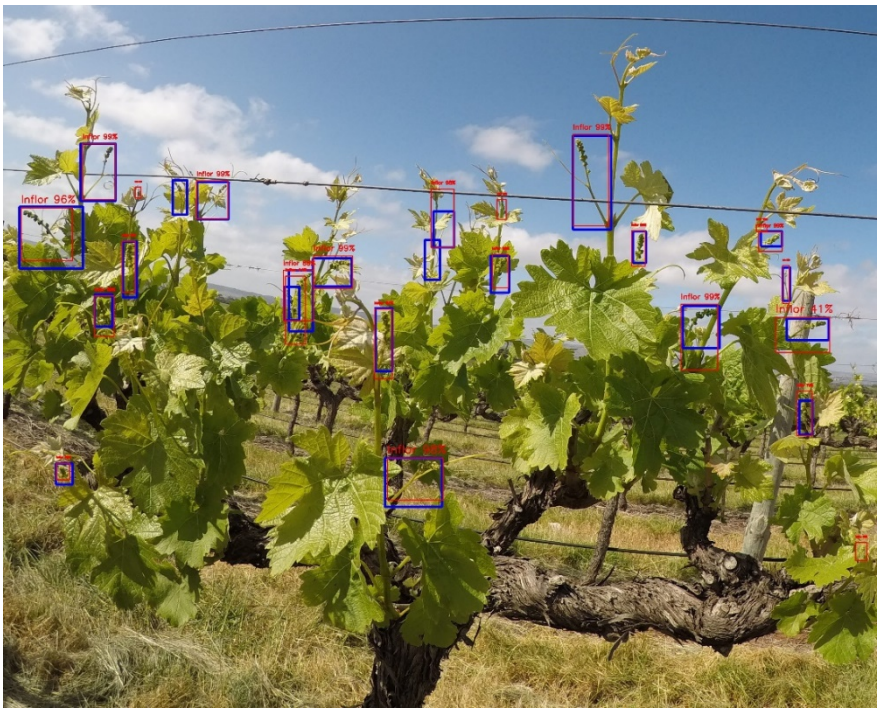


Figure 12 – cropped image section demonstrating the correspondence between hand and machine labelled inflorescence bounding boxes.

were to be used to estimate bunch mass (see below) and heavy occlusion was likely to make this difficult from RGB images.

To assist with these difficulties, a classification system was added to the manual labelling, with “full”, “partial” and “hidden” classes. The full class representing exposed bunches, partial representing bunches where only a few berries were visible and hidden representing potentially perceived fruit where it was uncertain whether fruit were there, primarily in dark shaded areas of the image.

The classes would then allow all visible bunches to be counted, possible bunches to be incorporated or not depending on accuracy in the rest of the system and heavily occluded bunches to be counted, but not used for bunch size estimation. Effectively, using classes both offered the opportunity to improve the bunch count estimate and to incorporate the bunch counting more easily into a full pipeline that counted bunches and estimate their mass.

In total 445 images were labelled for each time-point, with the three classes used in both sets of images. Figures 13 and 14 represent examples of the hand-labelled pre-veraison images and the detector output respectively, whereas Figures 15 and 16 represent the hand-labelled and detected pre-harvest bunches respectively.

The precision and recall of the two detectors trained were lower than for the inflorescence detector, likely due to the smaller number of input images and more difficult task due to varying occlusion and the use of classes, but of a similar magnitude (Tables 2 & 3; for pre-veraison and pre-harvest respectively).



Figure 13 - examples of hand labelled video frames demonstrating the range of conditions experienced during pre-veraison bunch imaging, including clear sky, sun in view causing glare, low light and difficult background objects such as trees or adjacent vine rows. Bounding box colours represent class used, green = “full” bunch; blue = “partial” bunch and red = “hidden” bunch.



Figure 14 - video frames from Figure 12 with bounding boxes determined through the deep learning based inflorescence detector. Colours as for Figure 12.



Figure 15 - examples of hand labelled video frames demonstrating the range of conditions experienced during pre-harvest bunch imaging, including clear sky, sun in view causing glare, low light and difficult background objects such as trees or adjacent vine rows. Bounding box colours represent class used, green = “full” bunch; blue = “partial” bunch and red = “hidden” bunch.



Figure 16 - video frames from Figure 14 with bounding boxes determined through the deep learning based inflorescence detector. Colours as for Figure 12.

Table 2 - sample output of pre-veraison bunch detector training, together with Precision – Recall results. Selected number of training epochs highlighted in bold.

No. of Epochs	**No. of Detections (GT = 1,533)**	**Recall**	**Precision**
2	TPs: 1,275, FPs: 584	0.832	0.753
5	TPs: 1,321, FPs: 897	0.862	0.743
10	TPs: 1,236, FPs: 354	0.806	0.755
12	**TPs: 1,225, FPs: 321**	**0.799**	**0.747**

Table 3 - sample output of pre-harvest bunch detector training, together with Precision – Recall results. Selected number of training epochs highlighted in bold.

No. of Epochs	**No. of Detections (GT = 1,247)**	**Recall**	**Precision**
6	TPs: 1,052, FPs: 600	0.844	0.756
7	TPs: 1,031, FPs: 442	0.827	0.758
14	TPs: 998, FPs: 327	0.800	0.735
15	**TPs: 1,003, FPs: 335**	**0.804**	**0.738**

Moving from digital ground truth to field estimates: the results above compare the number of objects detected by a human in a video frame with those detected by the machine learned detector. However, although still images can be used to make spot in-field estimates, the aim of the work and the reason for videoing entire rows was to count at scale, whole rows or whole vineyards. As noted above, even a perfect detector will not result in accurate estimates in the vineyard unless each inflorescence is only seen and counted once. In the videos, each inflorescence was seen in many frames from slightly different angles. Indeed, this was beneficial as it gave multiple opportunities for an inflorescence to be seen and counted, even where it might be occluded from one viewpoint. Two methods were trialled to account for this, tracking and mosaicking.

The initial tracking algorithm used a hard-coded number of maximum frames for which the object could be missing before being counted as new. This was a compromise between dealing with occlusions or glare over multiple frames and the trajectory estimate failing due to increasing inaccuracy as the number of frames between detections increased. Typically, the error between whole row counts of inflorescences (the object on which the tracking system was initially built) and hand counts (not hand labelled digital ground truth) of inflorescences for the row was about 20% (data not shown). Replacing the hard-coded number with an additional machine learning step reduced this significantly, to error rates of approximately 5% for inflorescences (Table 4). A still from a demonstration video, with inflorescences being counted, is provided in Figure 17.

Table 4 – detected inflorescences by combined tracking-detector pipeline compared to ground truth hand counts for sample rows from two vineyards.

Rows	Ground-truth counts	Proposed method	Error
R83-east	2,628	2,742	4.3%
R85-east	2,516	2,390	-5.0%
R86-west	2,241	2,109	-5.9%
R87-west	2,337	2,288	-2.1%
R97-north	5,917	5,598	-5.3%



Figure 17 – frame from video taken at E-L 14 growth satge at the Adelaide Hills vineyard with detected inflorescences and their count overlaid. Note the complex background as a result of other vinyard vegetation.

The development of the tracking algorithm was undertaken using the inflorescence videos and detector, but it was also applied to the videos for bunch counting. The error rates were higher than for the inflorescences, but the tracking still needs to be further tuned for use with bunches and, although part of the purpose of the tracking is to alleviate imperfect results from the detectors, the detectors for the fruit had a higher error rate than the inflorescences (Table 5).

The mosaicking approach worked by taking only a small portion of each video frame and producing a mosaic where each part of the row only appeared once. Figure 18 provides a thirteen panel section of row and a close up of three panels, both at reduced resolution. Due to background artefacts generated as part of the mosaicking process a detector specifically for the mosaic images had to be trained on newly labelled images (see also Figure 20). This was conducted as a proof of concept with a very limited number of labelled ground truth images and was successful, albeit with reduced precision and recall compared to the results above (data not

shown). It is likely that a larger effort to label images and optimise the detector would improve this.

Table 5 - detected pre-veraison and pre-harvest bunches by combined tracking-detector pipeline compared to ground truth hand counts for sample rows from two vineyards

Rows	Pre-veraison fruit		Pre-harvest fruit	
	Ground-truth counts	Proposed method	Ground-truth counts	Proposed method
R83-east	855	1337	1071	1664
R85-east	948	1246	945	1264
R85-west	814	1049	1029	1345
R86-west	798	672	770	945
R96-south	4253	2164	4538	3107

One other effect of the mosaicking was that it made detection of posts (and potentially trunks) relatively straightforward to detect (see Figure 18), this could assist in generating spatial information from the data.

When considering that all the data used for this work was from commercial vineyards, no special equipment was used and view angles included glare from the sun on sunny days, it is clear that a combination of detector and tracking algorithm offers a real opportunity to estimate inflorescence and bunch number in commercial vineyards at scale, with multiple rows or whole blocks able to be quickly imaged. The processing is computationally intensive, but could be implemented on recent consumer devices based on GPUs. The mosaicking application is also viable, but offers less resilience to partially occluded objects without many mosaics being generated at different view angles and tracking of objects between them.

Bunch size estimation: methods for estimating bunch size will be further explored in follow-up project CSA1701-3.1, but stereo imaging was trialled during this project as a means of determining bunch volume. Stereo imaging allows a disparity map to be produced from the image pair, which allows the depth from camera to be inferred (Figure 19a-c). This in turn can be used to build a 3D mesh of the visible parts of the bunch (Figure 19d), based on colour segmentation (see below).

Red fruit contrast strongly with the other components of a vineyard, particularly the canopy amongst which they are normally imaged. This makes them particularly suitable for colour segmentation, where the colour of an object is used to select it in an RGB image. This is commonly used to identify mature grapes in RGB imaging (e.g. Liu & Whitty, 2015). To establish a basic process for use within our project we used colour space transformation and colour ratios to identify the objects with colours equating to red grapes and mathematical morphology techniques to filter out small and unsuitable shapes. This process is not computationally



Figure 18 – reduced resolution mosaicked of vineyard row with posts and panel detected. Large section of row (top) and three panel close-up (bottom).

intensive and can be used on any sort of RGB image, whether stereo (Figure 18), video frame (Figure 19 top) or mosaicked image (Figure 19 bottom).

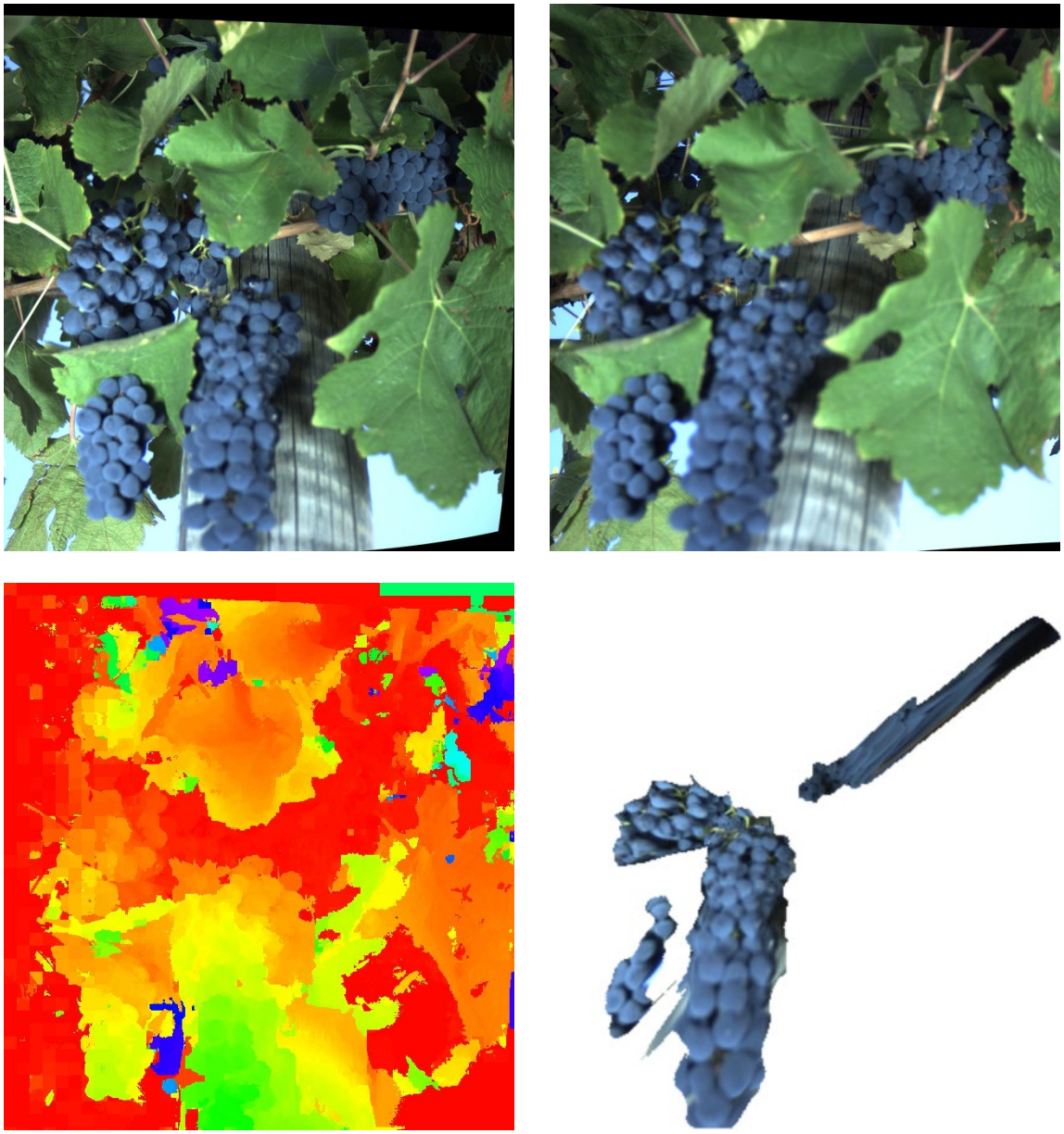


Figure 19 – example of a stereo image pair (top), disparity map (bottom left) and extracted 3D bunch using colour segmentation and depth information from the disparity map.



Figure 20 – examples of colour segmentation of red grape bunches i) a machine vision camera images with fruit present (top left), ii) without fruit present (top right) and iii) an image mosaic created from RGB video.

Bunch mass with FoPEN: electromagnetic modelling of berries, leaves and stems indicated that at frequencies above 7 GHz leaves could mask the grapes and that stems caused interference in the 5-6 GHz range (Figure 21 top). Not surprisingly, there was a high signal from water at 2.4 GHz. Modelling of an individual berry suggested peaks at 4 and 5.5 GHz (Figure 21 bottom).

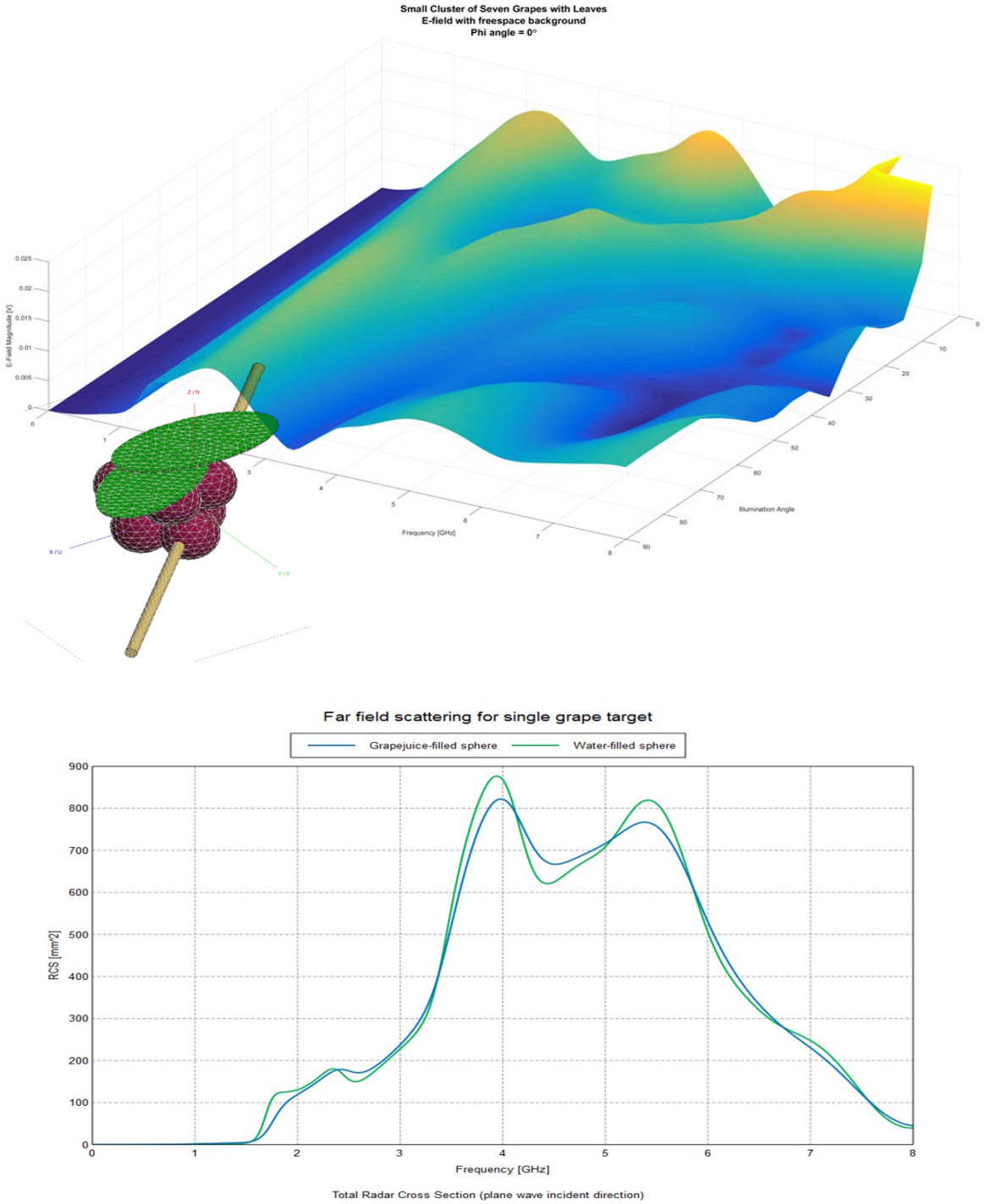


Figure 21 – e-field magnitude of modelled mixed grapevine sample from 0-8 GHz frequency and 0-90° illumination angle (top) and far field scattering for a single grape berry (bottom).

The initial testing using the rack system looked at the sensitivity of the hardware being used. With no foliage, group of five grapes was clearly visible at 200 mm distance from the sensor and there was some signal through to 450 mm distance. With light foliage in place at 450 mm from the sensor, a group of ten grapes was apparent in the signal at 650 mm (Figure 22). Again, there was some signal at a greater distance (800 mm), but it was less clear. These results gave an indication of the sensitivity of the radar system and suggested that very small bunches (<20-40 g) might not be detected in a vineyard situation with heavier canopies.

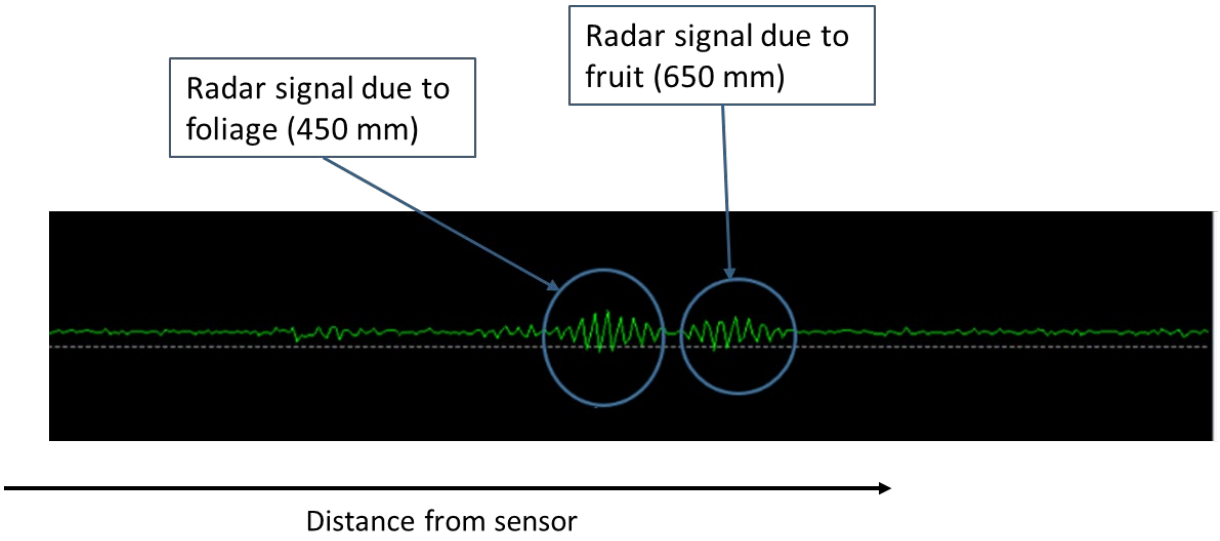


Figure 22 – single radar return of grape berries occluded by leaves on the test rig.

A number of field tests were made during the 2017/18 and 2018/19 seasons, including trials with different radar sensors (data not shown). Most of these evaluation campaigns consisted in scanning part rows with different arrays and array positions on the vehicle, then removing sections of leaves and/or fruit. For example, Figure 23 provides a demonstration with a single return, where there is signal from canopy and fruit, then with the canopy removed and finally with the fruit removed, leaving the background signal from the vineyard structure.

A software tool was developed to visualise the complex signal as the radar sensor moved down a row and quantification of fruit would of course require the analysis of adjacent returns. Figure 24 demonstrates continuous data with the same foliage+fruit, no foliage, no fruit situation. The complexity of the data is apparent, with amplitude (z axis), time/distance from sensor (y-axis) and distance down the row (x-axis). In both figures the noise at the start of the data (adjacent to the sensor) is generated by the sensor itself and not part of the path through the vine canopy.

For the 2019/20 season a different array was used, with multiple sensors angles to the vine, to test whether the consistent radar cross-section of a bunch (approximately round) was more easily discernible from leaves, as no leaf could provide the same cross-section at all angles due to being flat. At the time of writing that data has not been fully analysed and will be assessed in the follow-on project CSA1701-3.1.

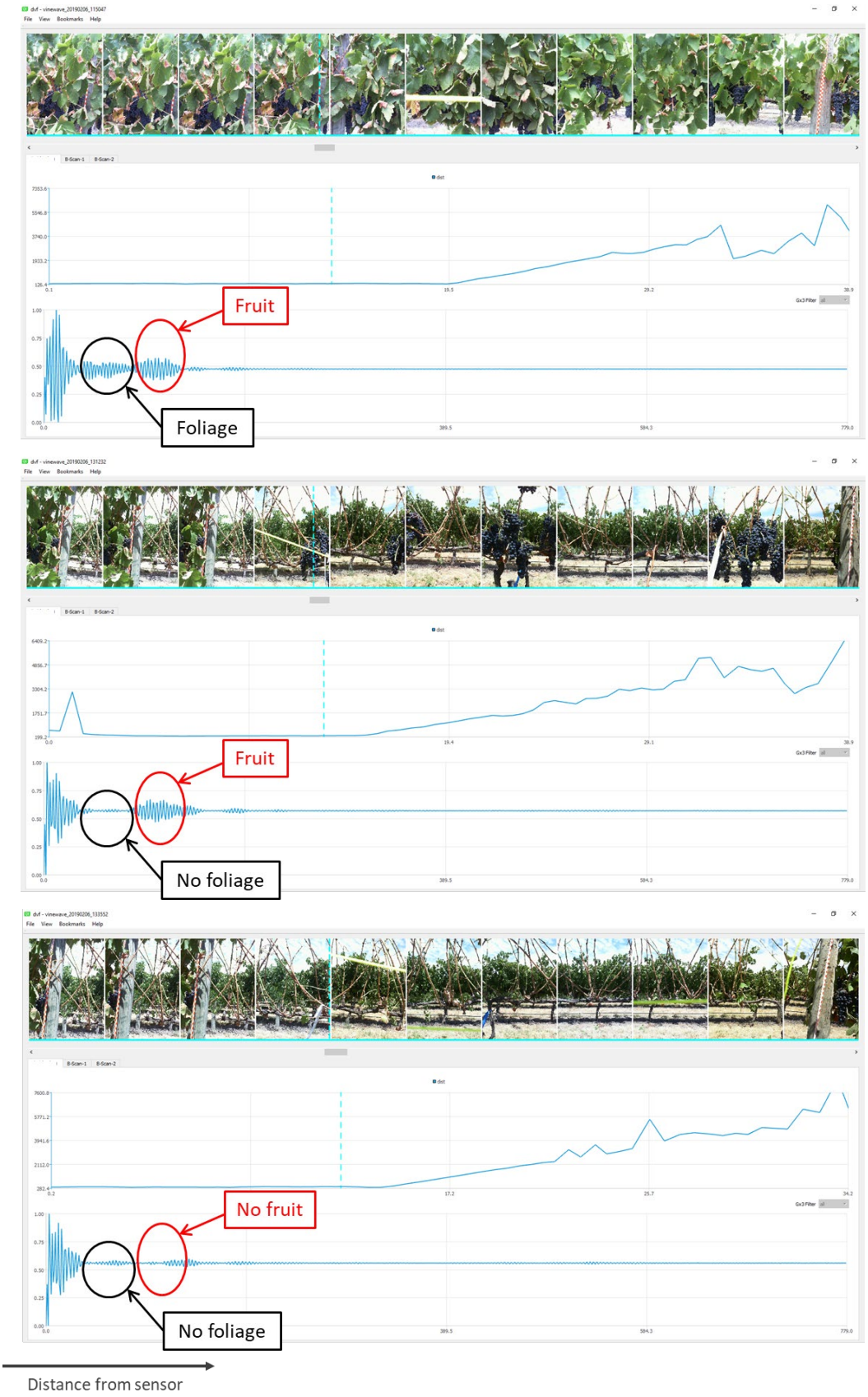


Figure 23 – screenshots from the radar software tool showing single returns and corresponding RGB images for intact vines (top), with canopy removed (middle) and with canopy and fruit removed (bottom).

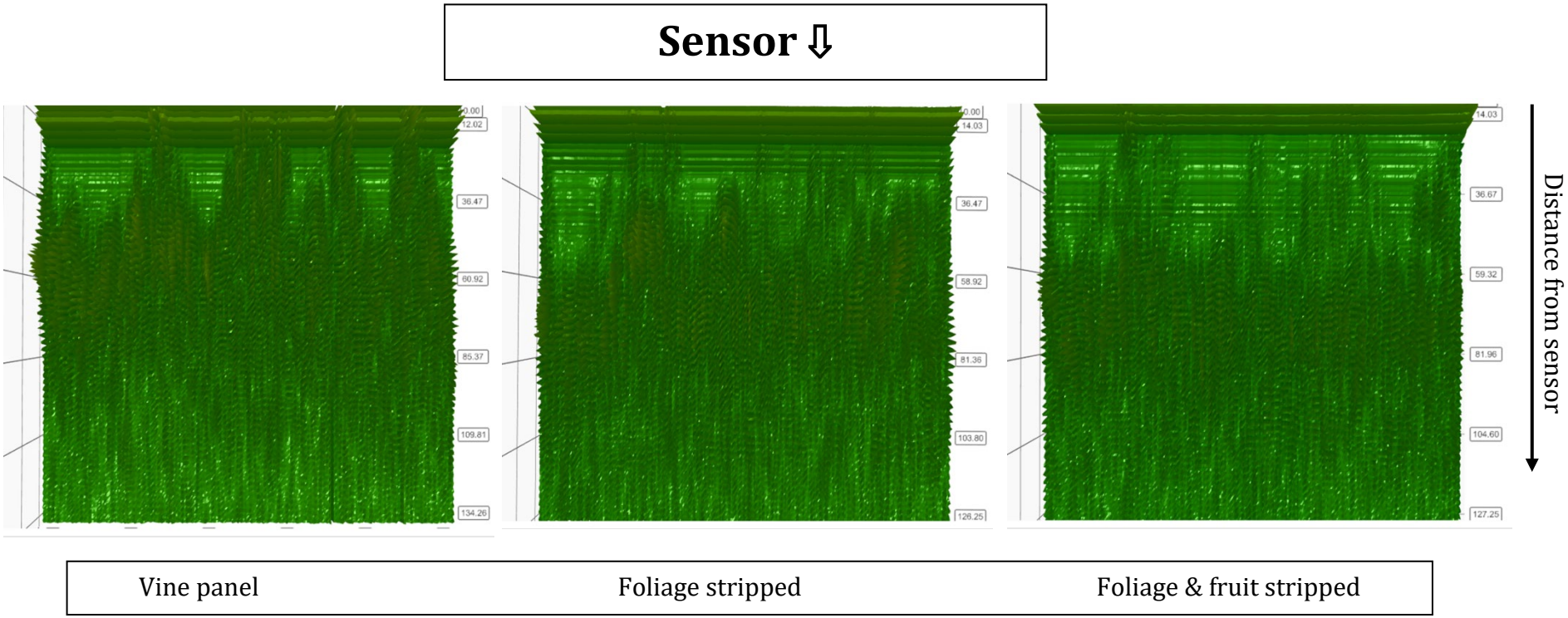


Figure 24 – radar returns from section of intact vine row (left), with canopy removed (centre) and with canopy and fruit removed (right).

3.1.2 Non-destructive measurement of fruit composition

Of the five data sets of lab-based fruit composition measurements collected, four were complete and incorporated into hyperspectral based models at the time of writing. A range of multiple-regression and machine learning tools were tested on these data sets, with a focus on Partial Least Squares (PLS), Least Absolute Shrinkage and Selection Operator (LASSO) and Random Forests (RF).

The initial models were built against data set A (red glasshouse grown berries, n=240), data set B (white glasshouse grown berries, n=720) and the two combined. Ten parameters from the Oenofoss instrument, plus total soluble solids (TSS) using a refractomer were modelled using the three tools and assessed through the coefficient of determination (r^2) between the predicted and ground-truth values for each berry. All three models produced reasonable results for both the individual data sets and the combined model (r^2 typically > 0.6 for data set A, 0.75 for data set B and > 0.8 for the two combined). Overall performance of the combined model was better than for white berries, which was better than for red berries, likely due to the increased number of replicates included. RF performed slightly better than LASSO across the 11 parameters, which had similar performance to PLS. There was considerable variation between parameters, however, and RF generally outperformed the other models on the more poorly predicted parameters, possibly due to a better tolerance for non-linear data (data not shown).

Data sets C (Shiraz, n=1440) and D (Sauvignon Blanc, n=1080) were larger again and consisted of field collected fruit. In both cases models trained on the individual data sets performed better than those on datasets A & B or on the combined A & B model, which had a similar sample number as dataset D. Figure 25 provides a range of individual berry spectra for data sets A-C. The difference between white and red berries can be clearly seen at the red wavebands (< 650 nm), but it is not clear that overall the field berries (C) were more clearly separated from the glasshouse berries (A & B) or that the red berries (A & C) were clearly separated from the white berries (B).

A combined LASSO model of datasets A-D inclusive, performed slightly less well than C or D modelled independently (r^2 was lower for most parameters), but better than A or B (data not shown). For yeast assimilable nitrogen (YAN), amino-N and tartaric acid, the multi-genic model was significantly poorer performing than models for the individual genotypes. Interestingly, ammonia-N remained well predicted. Table 6 provides the performance of the multi-genic model overall and for each dataset (genotype), together with the relative performance vs the independent models. The fifth dataset collected encompassed 1800 berries from four field grown varieties and will be incorporated into the combined model through CSA1701-3.1.

The datasets cover the entirety of maturation, from pea-size to post-harvest, allowing fruit maturation to be tracked in more detail (composition) from early in the season than currently occurs. Figure 25 provides a visual display of berries and bunches through maturation and gives an indication of the variation in the pixel values for a single berry and across a bunch.

Application of these models depends not just on the goodness of fit, but also the amount of variance around the regression. A data set with significant noise around the prediction can still result in a high r^2 value. The potential effect of this can be seen in Figure 28, where all parameters have a good r^2 (>0.9), but differ in terms of the variation (error) around the prediction. For example, berry sugar (predicted as both TSS and Baumé) has a very tight

relationship (low RMSE), meaning that TSS can be reliably predicted to within approximately 1°Brix across a very wide range (5 to 28°Brix) and possibly this could be improved further by excluding the poorly predicted data at low TSS (<2°Brix) from the modelling process. In contrast the models for pH and TA still have a very good r^2, but much more error around the predicted value. This is likely due to the concentration of these datasets in certain ranges and will be improved with additional data.

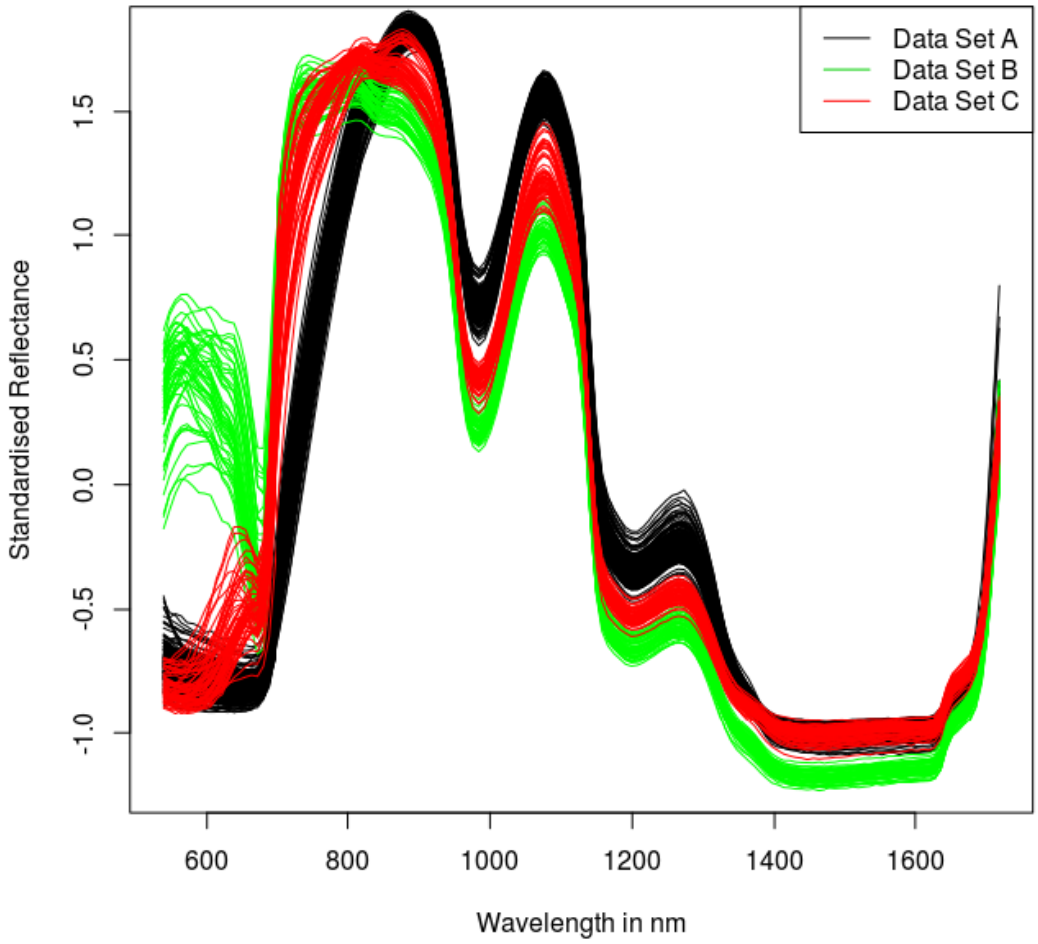


Figure 25 – normalised berry reflectance spectra for three datasets, red microvine (A), white microvine (B) and field-grown Shiraz (C).

Table 6 – coefficient of determination vales for model prediction vs lab measured berry composition for the multi-genic model across all data and per dataset, with the %change in the r^2 fit compared with models trained to the individual dataset.

	Multi-genic	**S. Blanc (field) (n=1080)**		**Shiraz (field) (n=1440)**		**MV red (GH) (n=240)**		**MV White (GH) (n=720)**	
	r^2	**r^2**	**% change**	**r^2**	**% change**	**r^2**	**% change**	**r^2**	**% change**
pH	0.94	0.94	-2.1	0.97	-3.1	0.36	23.7	0.84	14.6
Total Acid	0.91	0.90	-2.2	0.96	-7.1	0.76	0.0	0.83	2.2
Malic Acid	0.92	0.92	-2.1	0.96	-6.1	0.50	-1.1	0.87	5.7
Tartaric Acid	0.60	0.38	0.0	0.67	-25.9	0.39	11.1	0.51	-13.0
TSS	0.98	0.98	-1.0	0.99	-1.0	0.92	2.1	0.97	0.0
Volatile Acid	0.83	0.79	0.0	0.88	-10.8	0.42	56.6	0.74	7.8
YAN	0.47	0.36	-11.3	0.53	-35.6	0.11	42.4	0.27	-14.5
amino-N	0.38	0.40	-40.6	0.34	-41.5	0.27	-25.5	0.55	-45.7
Gluconic Acid	0.86	0.87	0.0	0.92	-9.5	0.52	48.3	0.84	1.2
Ammonia-N	0.78	0.76	-3.7	0.79	0.0	0.27	21.9	0.69	1.3

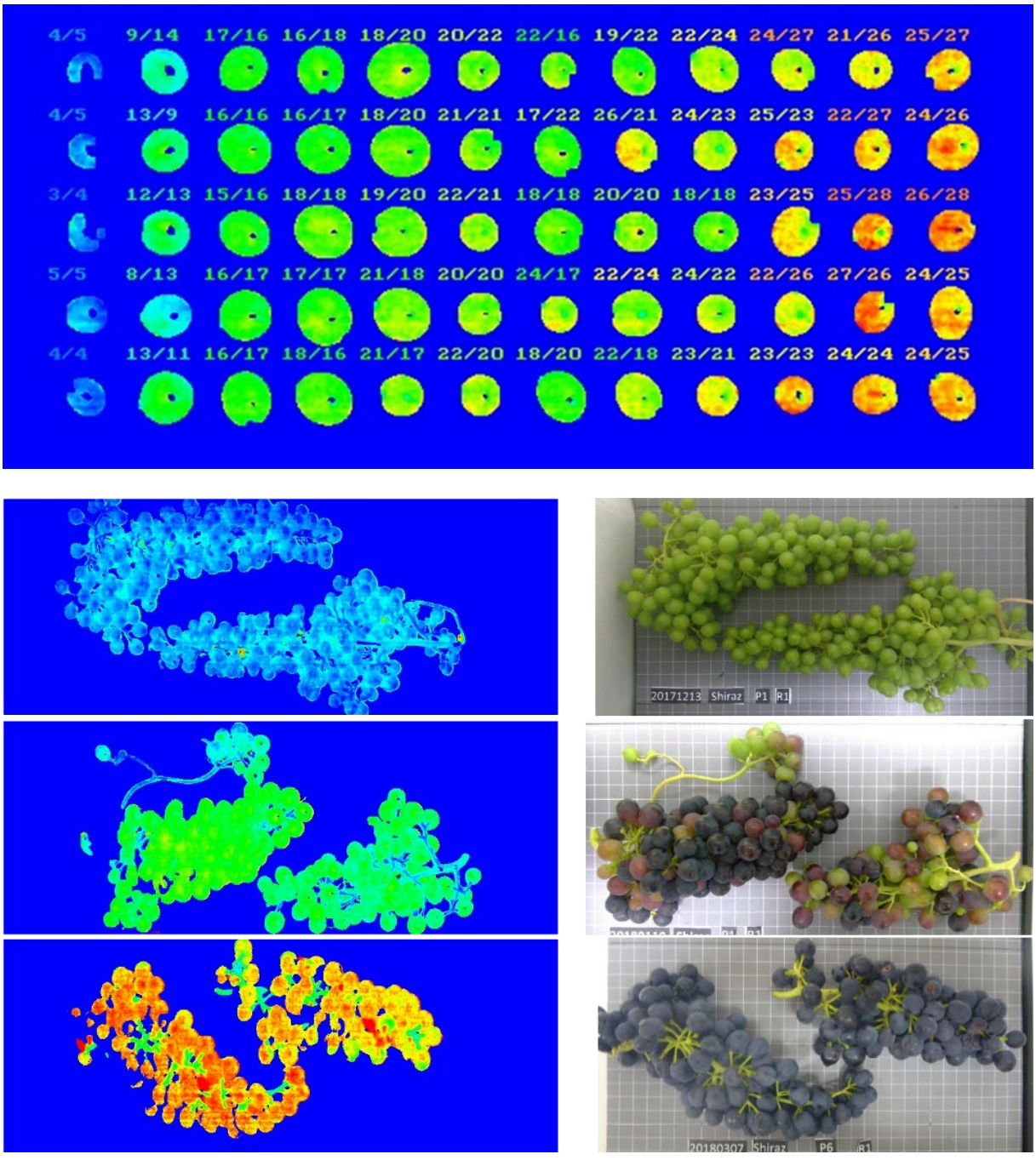


Figure 26 – visualisations of berry TSS for individual berries (top; numbers are predicted value / measured value) and whole bunches (bottom) at different stages of maturity.

The hyperspectral calibrations for berry composition demonstrate the potential for non-contact assessment without the need for laboratory analysis, but additional steps are required before they can be used at scale in the field. Firstly, either a suitable light source needs to be used with the sensor, or the calibration needs to work with reflected sunlight. Secondly, the object of interest (grape) needs to be separated from the background and surrounding vine material in the data. Thirdly, the data collection needs to be scaled to the field and the adequate representation of the collected, necessarily lower resolution, data demonstrated.

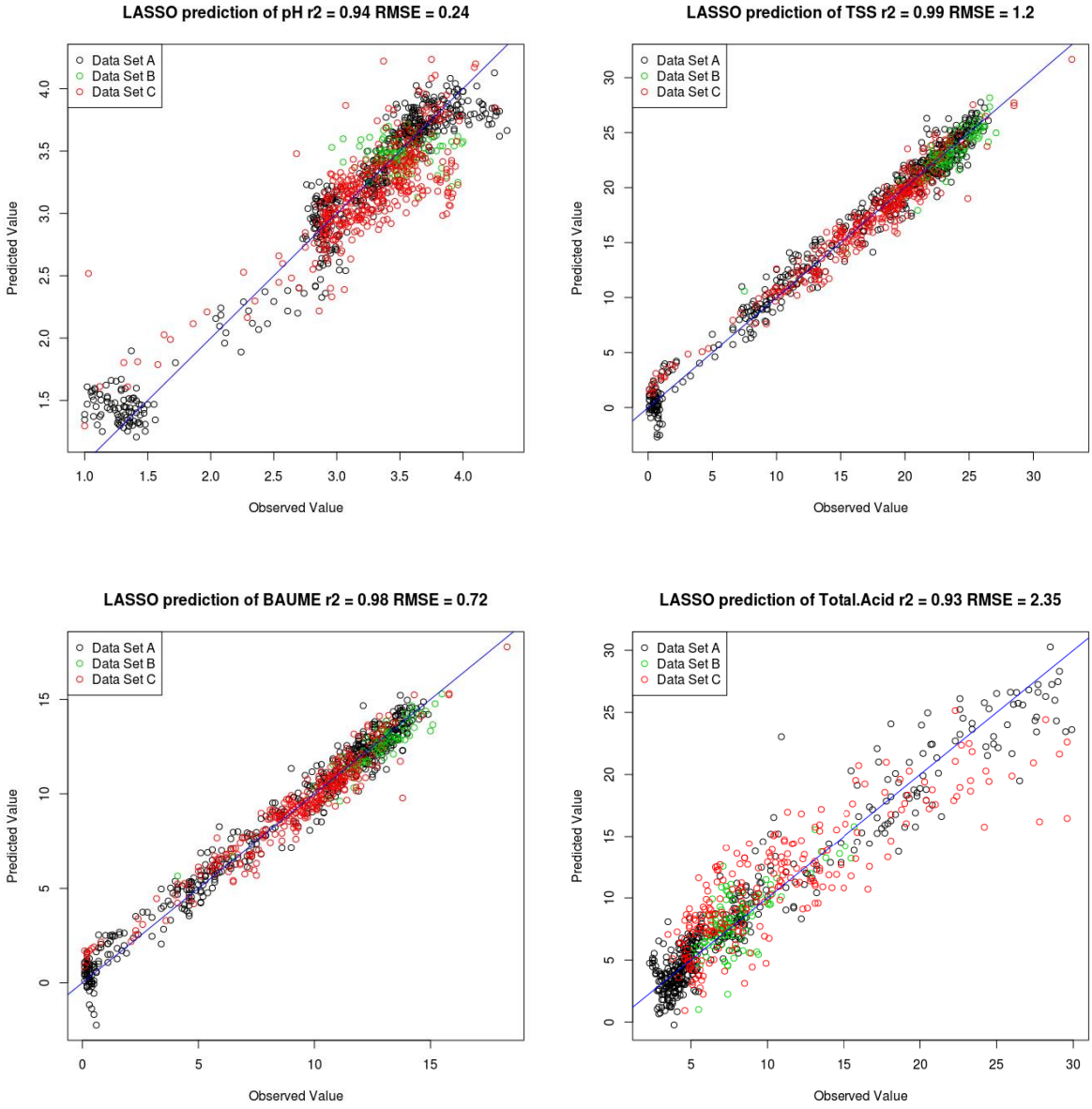


Figure 27 – predicted vs observed results for the combined dataset A,B,C model for berry composition paramters. TSS is measured using a refractometer, whereas Baumé is from the Oenofoss instrument.

The first field trial with a hyperspectral camera was undertaken in 2016/17, prior to the main Headwall instrument being available, and use to test the impact of using solar radiation as well as the practicalities of the Ximea camera. As can be seen by the banding in Figure 28 and the reference target reflectance spectrum in Figure 29 top, there is significant non-linearity in the spectrum of solar radiation at the earth's surface. The use of a white reference allows this to be adjusted for, just as with the lab-based set-up, however, at wavelengths where there is very little solar radiation this is inadequate, leading to artefacts in the corrected reflectance spectrum when compared to full spectrum lighting (Figure 29 bottom left). An attempt to limit the impact of wavebands where solar radiation is high, requiring the camera integration time to be set to avoid saturation, meaning that little light is received in some other bands, using partial shading

with the Ximea camera tests was ineffective (Figure 28). It was clear that reliance on solar radiation would result in some bands being unable to be included in calibrations, despite reference surface corrections.

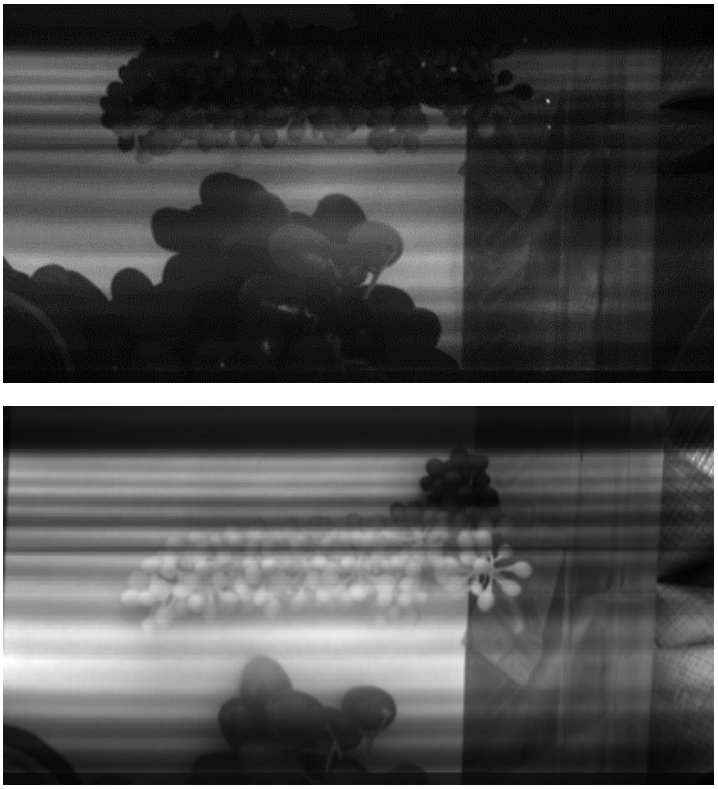


Figure 28 - sample images for a Shiraz bunch taken at pre-veraison, together with a table grape bunch under full sun and shade. The banding on the images is due to the different light available at the different wavebands due to the non-uniform nature of solar radiation. The images were not corrected to reflectance

Using a limited number of wavebands also offers the option of using a lower cost multi-spectral snapshot sensor for field imaging, such as has been trialled by other groups (Tang et al. 2019). Of the three analytical tools used for the hyperspectral calibrations, LASSO is particularly useful for determining the wavebands that drive the calibration as it is sparsity promoting, that is, it will apply a zero weight to inputs (wavebands) that have little predictive power. An example of the LASSO weights is given for the field grown Shiraz (dataset C) Baumé model and for the equivalent datasets A-C combined model in Figure 30. Many of the wavebands are indeed set to zero, providing a first pass in choice of bands for a multispectral camera. These results were then utilised to select seven wavebands per fruit composition parameter, according to the size of weight in the model, whether neighbouring wavebands had a similar weight (only one chosen if so) and to avoid features, such as positive and negative weights in adjacent bands, which would be difficult to achieve in a multispectral camera due to bandpass limitations of the filters employed.

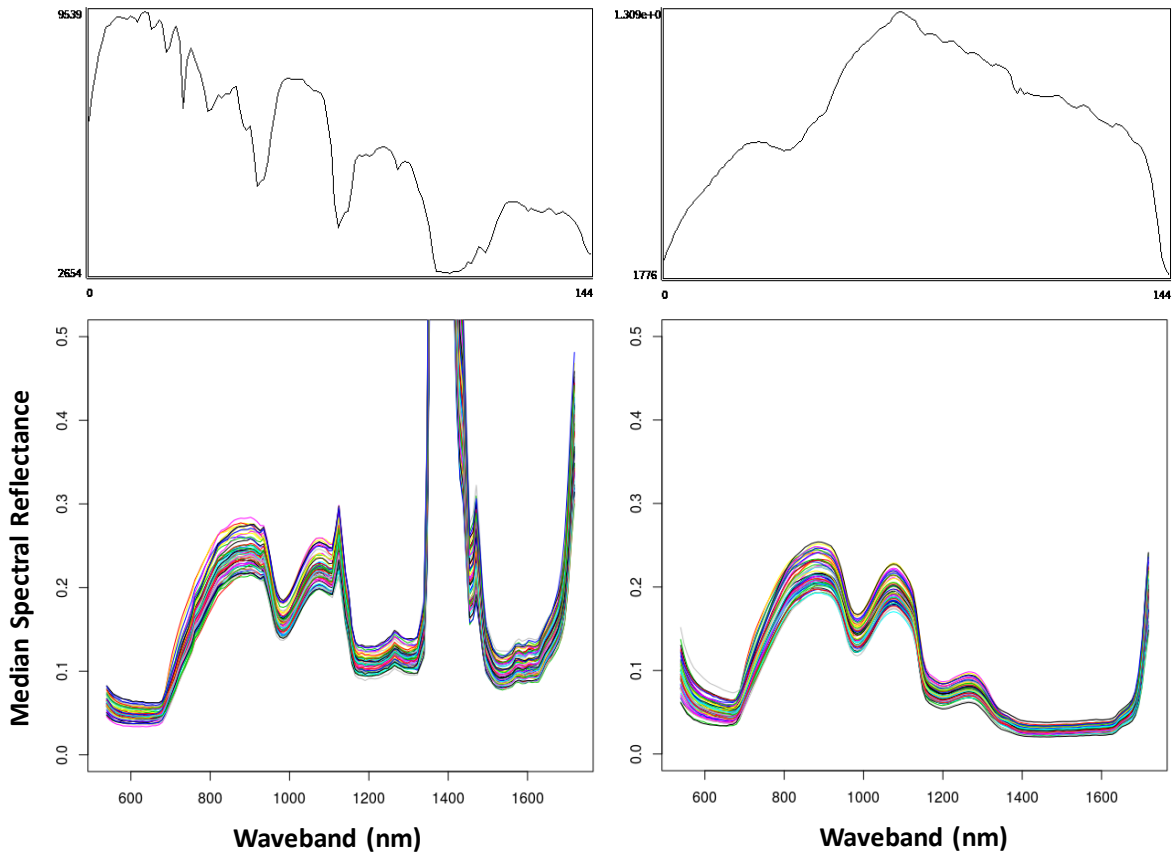


Figure 29 – illumination spectra (top) and corrected berry reflectance spectra (bottom) for sunlight (left) and halogen lighting (right).

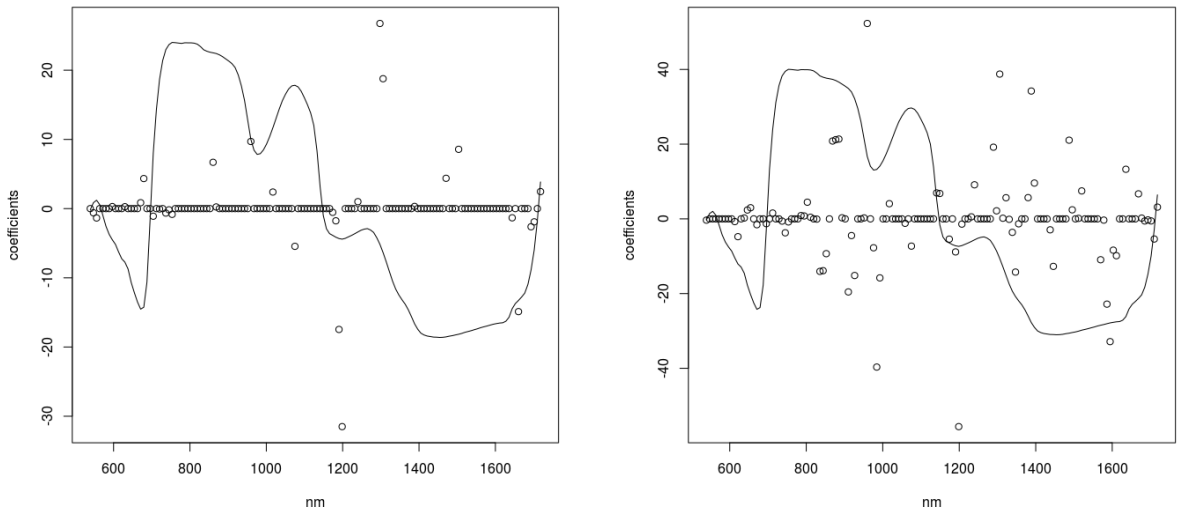


Figure 30 – per waveband weightings from LASSO models of berry spectra vs Baumé for a single dataset (C; left) and three combined datasets (A,B,C; right).

Both the r^2 and RMSE values suffered when the contributing wavebands were reduced to seven and there were large differences in the contributing wavebands between some of the

compositional parameters (Table 7). However, for berry sugars (TSS or Baumé) and pH the predictive power of the reduced model would likely be sufficient for commercial use, particularly when applied to many spectra across a vineyard row or block to obtain a median or mean. Some of the selected wavebands are in the dips in solar radiation and unlikely to be suitable without additional illumination, but the results suggest that a model could be built relatively simply to avoid the areas of the spectrum where solar radiation is at a minimum, particularly if the whole available spectrum was used rather than a multi-spectral approach.

Table 7 – comparison of coefficient of determination and RMSE for LASSO models including all wavebands and seven selected wavebands.

Measure	All wavebands r^2 (RMSE)	Select wavebands r^2 (RMSE)	Wavebands selected
pH	0.94 (0.24)	0.93 (0.25)	1207, 1306, 1355, 1397, 1446, 1495, 1660
Total Acid	0.93 (2.35)	0.84 (3.47)	704, 836, 893, 992, 1198, 1322, 1487
Malic Acid	0.92 (2.85)	0.86 (3.66)	662, 745, 885, 992, 1165, 1388, 1479
Tartaric Acid	0.74 (1.02)	0.61 (1.19)	580, 679, 918, 1198, 1306, 1545, 1643
TSS	0.99 (1.20)	0.96 (2.02)	868, 984, 1198, 1306, 1388, 1512, 1594
Baumé	0.98 (0.72)	0.96 (2.03)	836, 959, 1140, 1198, 1306, 1388, 1594

The second requirement for field application of hyperspectral calibrations was the ability to separate fruit from background. A reliable classification model to separate vine parts, including fruit, from each other and background objects was built as part of the sister project, CSA1601, and field-tested pre-harvest during the 2018/19 season. As the Headwall linescan camera was mounted on the field vehicle and there was no attempt to prevent vertical movement by the vehicle as it scanned down a section of row the resulting image was significantly distorted (Figure 31 left). Furthermore, because the xenon flash was used to illuminate the vines, scans could only be collected at 1 Hz (the camera has a capability to scan at a much higher rate), which meant that even when the vehicle was driven very slowly, only a few scans were collected across a bunch, giving a compressed visual appearance.

The classification model was used to select pixels which contained berry spectra and then filtered for connected components, thereby selecting groups of pixels which might be expected to be from a single bunch (Figure 31 right). The spectra from each prospective bunch (62 in total from this test run) were plotted and were visually similar to lab collected spectra for ripe Shiraz fruit. Fruit TSS was then predicted for the prospective bunches, using the lab-based model for TSS. The results were in the range expected for maturing fruit, but under-estimated compared with a small number of bunches that were used for a ground-truth TSS estimate in the laboratory. Lab testing of the xenon flash also typically underestimated TSS relative to the full range lab lighting system (data not shown).

Overall, this field-test clearly demonstrated the potential to undertake hyperspectral imaging in the field, use pre-built libraries to identify pixels that represented fruit and then predict fruit composition from pre-built calibration data.

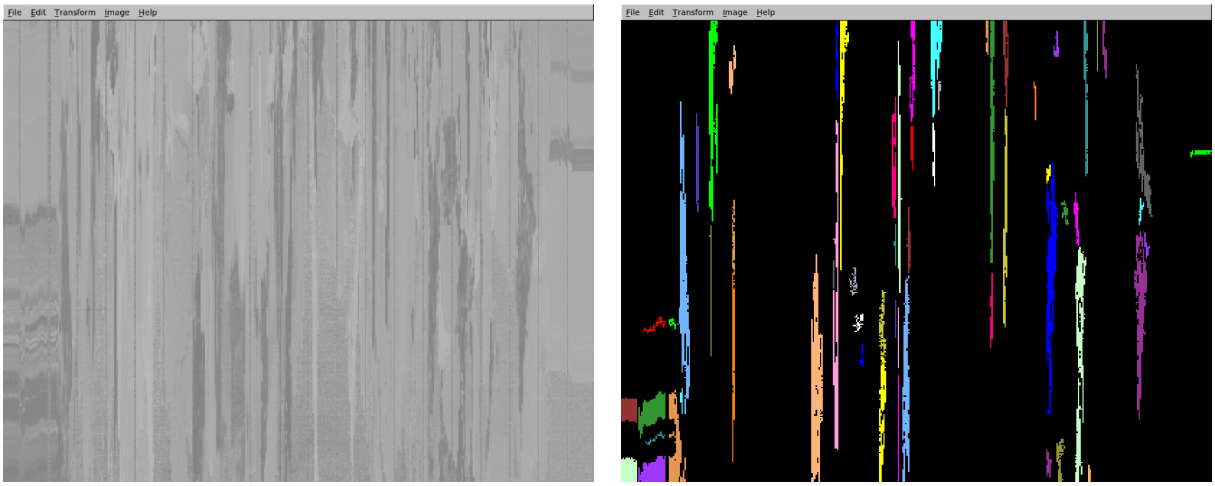


Figure 31 – field collected hyperspectral data (single waveband; left) and segmented for grape bunches (right).

Despite the success of this test, the project aimed to develop a commercially viable method for non-destructive, on-the-go assessment of fruit composition and a system based on an eVNIR camera such as the Headwall system used would likely cost >$200k, limiting uptake. As mentioned above, one approach is to use multi-spectral imaging, but it would be difficult to construct a camera with a realistic number of filters that could be used to separate fruit from background and produce predictions of multiple fruit components. Furthermore, a moving 2-D camera would require additional processing for frame overlap, even if data could be collected without blurring pixels.

As a result, a device was built using two point-spectrometers, which covered the same eVNIR range. This device was much cheaper (approx. 10% of the eVNIR linescan camera), but could not be used for imaging. However, together with a halogen light source, it could be driven around a vineyard collecting large volumes of spectra, which can be filtered for fruit in the same way as described above. Initial validation and testing occurred in the laboratory and field at the end of the 2019/20 season and the results will be finalised as part of the CSA1701-3.1 follow-up project.

In summary, eVNIR hyperspectral imaging has proven a highly suitable technique to assess fruit composition throughout berry maturation, even prior to veraison. We have explored a number of options for translating this from a lab-based to field-based technique and demonstrated the viability of these. Further refinements will provide the most cost-effective path forwards towards potential commercial use.

3.1.3 Non-destructive measurement of fruit condition

The final aim of the project was to test and develop technology to assess fruit condition in the vineyard. Visible disease assessment by hyperspectral imaging was clearly demonstrated in the sister project CSA1601 and disease assessment at winery intake by hyperspectral imaging was being examined through Wine Australia project AWR1601 (Dambergs et a. 2019l). Consequently, the hyperspectral work was focused on fruit composition, with berry condition limited to undertaking testing for shrivel in red grapes.

As with intact berries, a large number of shrivelled Shiraz berries were scanned with the lab-based HSI system and the resulting data processed in the same way. When intact and shrivelled berry spectra were compared there were very clear differences in spectral shape (Figure 32), which were used to build a classifier.

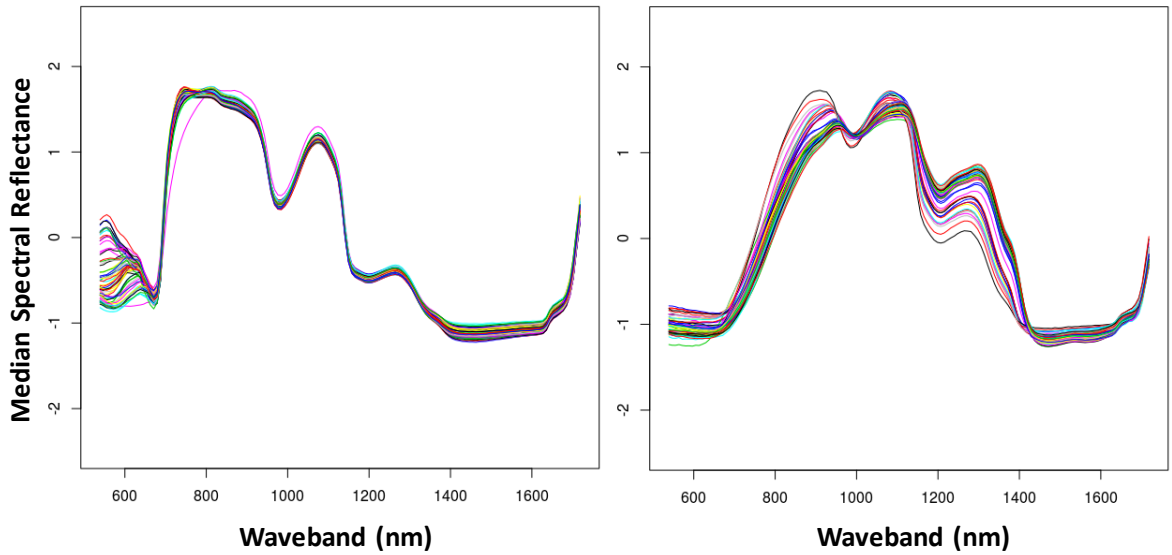


Figure 32 – normalised spectra for undamaged Shiraz berries (left) and shrivelled Shiraz berries (right).

The classifier was used with whole field grown bunches that were manually scored for proportion of shrivel, producing an estimate of the proportion of fruit pixels that were shrivelled berries (e.g. Figure 33). Testing against manually scored bunches revealed a broad agreement between the HSI classifier and the ground-truth (Table 8), noting that as the scoring was done by visual estimate, it is quite possible that the HSI classifier was the more accurate where estimates disagree to some extent.

Table 8 – Predicted percentage of shrivel in Shiraz bunches at five dates during late stage maturation vs visual scoring.

Date	100%	80%	60%	40%	20%
07/03/2018	99	93	72	49	27
28/02/2018	99	69	44	42	48
22/02/2018	97	83	79	48	51
14/02/2018	99	86	59	53	34
08/02/2018	98	72	72	42	42

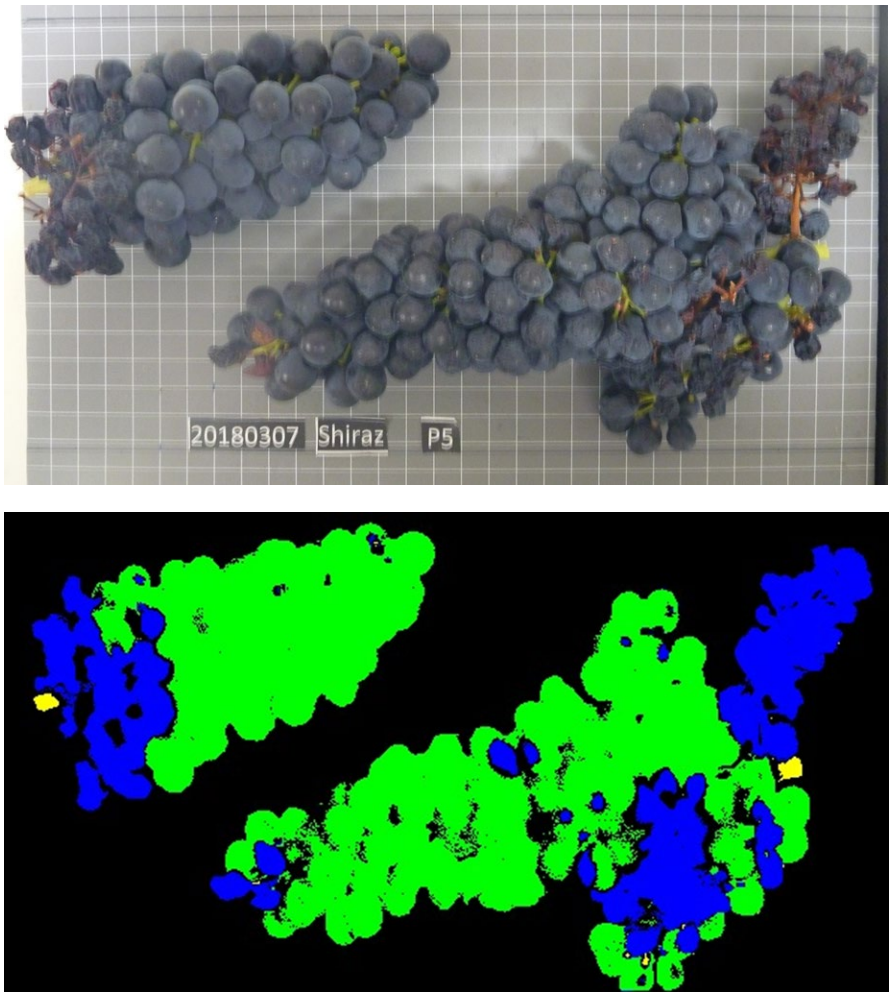


Figure 33 – RGB image (top) and per pixel classification from HSI scan for presence of shrivel.

3.2 Contribution to program objectives

The program objectives of program RnD4Profit-15-02-018, Digital Technologies for dynamic management of disease, stress and yield, were:

a) generating knowledge, technologies, products or processes that benefit primary producers;
b) strengthening pathways to extend the results of rural R&D, including understanding the barriers to adoption; and
c) establishing and fostering industry and research collaborations that form the basis for ongoing innovation and growth of Australian agriculture.

The project has focussed on developing, prototyping and assessing technologies and data analytic pipelines that will benefit the winegrape growers of Australia. Specifically, the project has:

- examined a range of methods and sensors to assess yield components for their potential use at commercial scale and developed data analysis pipelines to utilise these;

- built a large library of NIR spectral data with accompanying lab analysis for berry composition (>3,000 individual samples) and utilised this to produce calibrations for 10 fruit quality and maturation parameters;
- developed a prototype NIR dual point spectrometer sensor that can be used to map fruit composition and maturity in the field;

Whilst the tools and technologies utilised through this project are at various technology readiness levels, the project team has maintained a dialogue with the industry partners (Accolade Wine, Rymill Wine, Katnook Estate), the industry reference committee (including members from a number of wine industry companies) and has presented work at key industry events, such as the Australian Wine Industry Technical Conference and Exhibition. The team are also in active discussions with Wine Australia about pathways towards commercialisation of the most suitable technologies

Finally, the project collaborated with a number of industry partners (see above) as well as undertook informal discussions with the TeleVitis team at the University of Rioja, who are developing related technologies for the Spanish wine industry, about future interaction, particularly in the area of hyperspectral imaging of vineyards. These collaborations are aimed at supplementing Australian skills and know-how with relevant overseas experience to provide additional resources and thereby improve the efficiency of developing tools for Australian growers.

4 Collaboration

Although not formally listed in the original funding application, a discussion with Accolade Wine (Alex Sas, Warren Birchmore) to undertake work described in this report at their Mclaren Vale vineyards pre-dated the project. Further collaboration with commercial wine companies and grape growers was critical for field testing of equipment and methodologies in a commercial setting. In addition to the collaboration with Accolade, the project leads arranged to work with Wingara Wine Group (Chris Brodie) and Rymill Wine (Shannon Sutherland, Josh Clementson) in the Coonawarra. The Woodside vineyard managed by Adam Loveys provided access to some highly managed vines of alternative varieties, but was unfortunately lost in the 2019/20 season bushfires, prior to the end of the project.

It is important also to note that the project facilitated high levels of collaboration within CSIRO, with three different groups based in three different states coming together to generate novel hardware and data processing pipelines that would have been difficult to undertake within CSIRO Agriculture and Food alone.

The collaboration with Accolade will continue through the CSIRO-Wine Australia project CSA1702-3.1, which will undertake further development of the technologies described in this report and commercialisation suitability assessment. It is anticipated that the collaborations with Wingara Wine Group (through their Katnook Estate winery) and Rymill Wine will also continue.

5 Extension and adoption activities

This project was intended to develop commercially applicable sensors and data analytics for future commercialisation. As such extension activities to growers has not been practical for much of the project. That said, all of the development work has been carried out in collaboration with commercial growers (see Section 4, above). However, information has been provided to the industry through:

- A plenary presentation at the 17th AWITC, with a video of the presentation and a conference paper available to attendees.
- A YouTube video of a presentation on the project to the Televitis group at the University of Rioja is available at: https://youtu.be/PsKBD46py_s

In addition, conversations with potential commercial partners have been undertaken in conjunction with Wine Australia.

6 Lessons learnt

As is often the case, contracting was slow, which delayed the start-up of the project and meant that much of the first growing season was missed.

This project brought together previously disparate groups within CSIRO and a number of wine companies with sites across South Australia. Whilst we consider this to have been a highly successful endeavour, the time required to ensure that all parties had a common understanding was more than expected and a greater allowance in the project planning for this may have eased the project start-up.

Whilst CSIRO is uniquely placed, in Australia, to bring together diverse teams from different research areas, it was still necessary to bring mechatronics and coding expertise into the domain knowledge team at Agriculture & Food as part of the technical support. This was only partially anticipated and probably delayed the optimal team structure being established.

7 Appendix - additional project information

7.1 Project, media and communications material and intellectual property

Peer reviewed papers:

No research papers for publication in peer reviewed journals have been published at the time of writing.

Conference presentations:

Digitising the vineyard: developing new technologies for viticulture in Australia. E.J. Edwards, M.R. Thomas, S. Gensemer, P. Moghadam, T. Lowe, D. Wang, R. Lagerstrom, C. Hargrave, J. Ralston. (2019) Poster presentation at the 21st GiESCO meeting, Thessaloniki.

Digital opportunities: from the vineyard to the winery door. E.J. Edwards, M.R. Thomas, S. Gensemer, P. Moghadam, T. Lowe, D. Wang, R. Lagerstrom, C. Hargrave, J. Ralston. (2019) Oral presentation and conference paper at the 17th Australian Wine Industry Technical Conference. Adelaide.

Intelligent systems for commercial application in perennial horticulture. Edwards, E. J. and Moghadam, P. (2020). Oral presentation at TropAg 2019, Brisbane. Abstract published in: *Multidisciplinary Digital Publishing Institute Proceedings*, **36**(1):59.

Datasets:

Khokher, Rizwan; Wang, Dadong; Edwards, Everard (2019): Viticulture Dataset: Grapevine Inflorescence Detection. v1. CSIRO. Data Collection. https://doi.org/10.25919/5de4546aeacce

Intellectual Property

IP produced by this project included:

1. Data analysis pipelines, based on applications of machine learning, for estimating inflorescence and bunch numbers from RGB video.
2. Datasets of hand-labelled images, taken from RGB video and prepared for use in machine learning tools.
3. New knowledge on the application of radar in the vineyard for estimating fruit mass.
4. Hyperspectral calibrations and datasets (hyperspectral images with matching ground-truth data) for winegrape berry composition and condition.

7.2 Equipment and assets

Headweall eVNIR Micro hyperspectral camera

Ximea NIR hyperspectral linescan camera

Hammatsu Point-spectrometers

7.3 Thermal Infrared Imaging – Initial Evaluations

Although not part of the original sensing suite, the results of an initial evaluation of Thermal-infrared (TIR) imaging are also described here for completeness.

A series of exploratory evaluations were undertaken to investigate the use of passive TIR imaging as a possible means to characterise fruit and vine properties as pre-cursors for yield prediction and estimation. TIR cameras are sensitive to wavelengths in the range of 1 µm to 14 µm, which is larger wavelengths than regular visible light cameras. A TIR camera measures black body radiation, which is a property that is directly related to an object's temperature. Thermal imaging is becoming increasingly affordable and there is a large body of work describing its potential use in vineyards to assist with irrigation management.

The underlying motivation for using a TIR camera was to assess the extent to which an increase in fruit metabolic activity during growth phases could be correlated with an increase in fruit temperature. A Xenix Gobi 640 GigE TIR sensor was used for both in-house and field evaluations.

Figure 34 compares visible and thermal infrared images of a grapevine with fruit mid-veraison, which was measured in controlled glasshouse conditions. For the thermal grey-scale image (right), the lighter colour represents object of higher temperature. A clear thermal differentiation was observed on mature fruit with respect to the hosting foliage.

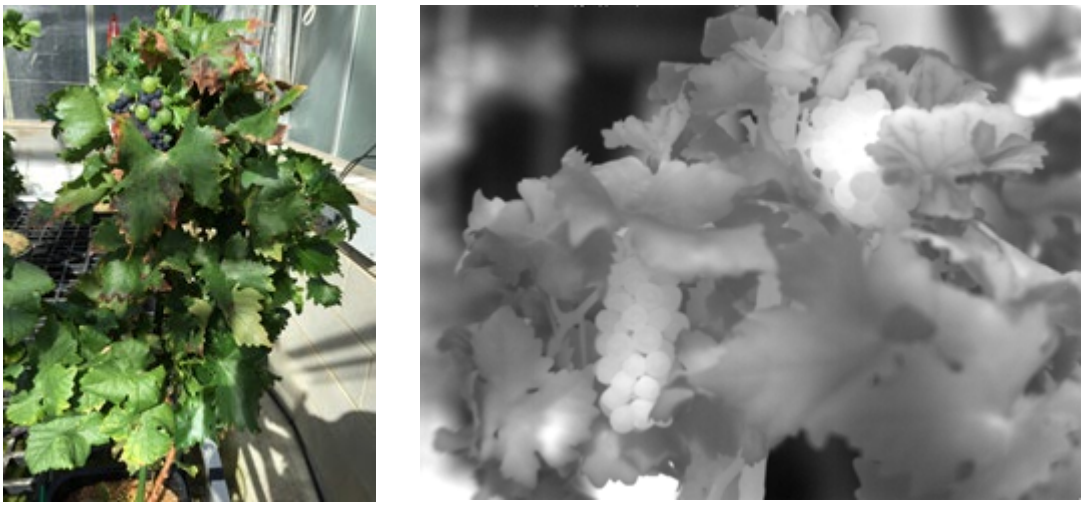


Figure 34 - Evaluation of thermal-infrared imaging of grapevine in controlled glass house conditions. A visible image of the vine with fruit at post-verasion stage (left) is compared with the thermal-infrared image (right).

Based on the positive results seen in the in-house glasshouse evaluations, vineyard-based measurements were then subsequently undertaken. For initial testing, fruit at various stages of development were hung to provide a variety of imaging targets and visible and thermal images acquired (Figure 35). Temperature variations of the fruit in both test and live configurations were noted. Although the temperature resolution of the particular Xenix instrument was not

sufficient for accurate thermal differentiation, the outcomes achieved through these initial results suggest that vineyard TIR has the potential to differentiate fruit from canopy in the vineyard, potentially also phenological stage, and offers the possibility of using a single instrument for assessment of vine water status and yield estimation.

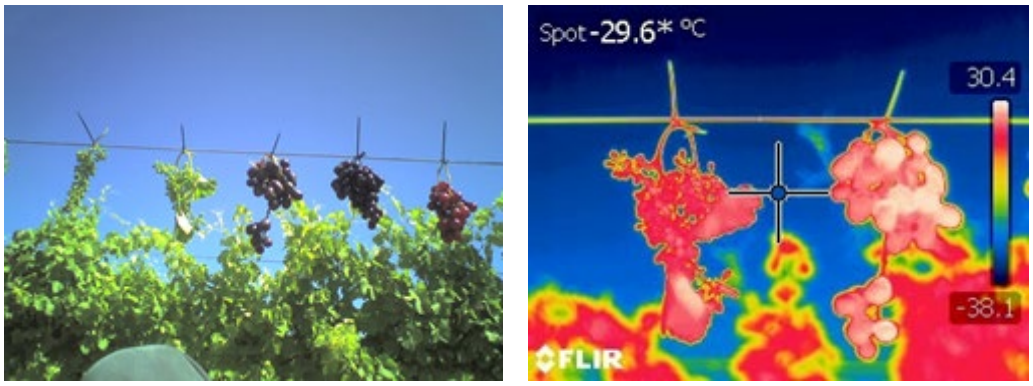


Figure 35 - Thermal-infrared imaging of grapevine in field conditions. Fruit at various stages of development are seen in visible (left) and thermal wavelengths (right).

7.4 References

Bankhead, P., Loughrey, M.B., Fernández, J.A. et al. (2017) QuPath: Open source software for digital pathology image analysis. *Sci Rep* **7**, 16878 https://doi.org/10.1038/s41598-017-17204-5.

Clingeleffer, P.R.; Sommer, K.J.; Krstic, M.P.; Small, G.; Welsh, M. (1997) Winegrape crop prediction and management. *Australian and New Zealand Wine Industry Journal* **12**: 354–359.

Clingeleffer, P.; Dunn, G.; Krstic, M.; Martin, S. (2001) Crop development, crop estimation and crop control to secure quality and production of major wine grape varieties: a national approach. Final report to Wine Australia.

Coombe, B. G. (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. https://doi.org/10.1111/j.1755-0238.1995.tb00086.x

Dambergs, B.; Jiang, W.; Nordestgaard, S.; Wilkes, E.; Petrie, P. (2019). Digital solutions for grape quality measures at the weighbridge. Final report to Wine Australia.

R. Gruna, K.-U. Vieth, M. Michelsburg, F. Puente León. “Hyperspectral Imaging - From Laboratory to In-line Food Sorting”. CIGR Workshop on Image Analysis in Agriculture, 26-27. August 2010, Budapest.

Iland, P.; Ewart, A.; Markides, A.; Sitters, J.; Bruer, N. (2000) Techniques for Chemical Analysis and Quality Monitoring during Winemaking. 120 pages. Patrick Iland Wine Promotions, Adelaide, Australia

M. Konstantinovic, S. Woeckel, P. Schulze Lammers, & J. Sachs. (2008). UWB Radar System for Yield Monitoring of Sugar Beet. *Transactions of the ASABE*, **51**(2), 753–761. https://doi.org/10.13031/2013.24372

Liu, S., & Whitty, M. (2015). Automatic grape bunch detection in vineyards with an SVM classifier. *Journal of Applied Logic*, **13**(4), 643–653. https://doi.org/10.1016/j.jal.2015.06.001

Mahlein, A., Oerke, E., Steiner, U. et al. (2012) Recent advances in sensing plant diseases for precision crop protection. *Eur J Plant Pathol* **133**, 197–209. https://doi.org/10.1007/s10658-011-9878-z

Nuske, S., Wilshusen, K., Achar, S., Yoder, L., Narasimhan, S., & Singh, S. (2014). Automated Visual Yield Estimation in Vineyards. *Journal of Field Robotics*, **31**(5), 837–860. https://doi.org/10.1002/rob.21541

Sun, C. (2002) Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques. *International Journal of Computer Vision*, **47**(1/2/3):99-117.

Sun, C. ; Beare, R.; Hilsenstein, V.; Jackway, P. T. (2006) Mosaicing of microscope images with global geometric and radiometric corrections. *Journal of Microscopy*, **224**(2):158-165.

Tang, J.; Petrie, P.; Whitty, M. (2019) Using Multispectral Imaging to Predict Grape Maturity. 17th AWITC, 21-24th July, Adelaide, Australia.

7.5 Staff

Ordered by role, research group and site

Research scientists

Everard Edwards
Mark Thomas
Stephen Gensemer
Giacomo Betti
Ryan Lagerstrom
Muhammad Rizwan Khokher
Qiyu Liao
Maoying Qiao
Changming Sun
Dadong Wang
Chad Hargrave
Jonathon Ralston

Technical staff

Don MacKenzie
Adam Smith
Najid Periera-Ishak
Paul Adkins
Wytse Gensemer
Mahmood Hassan

Paul McPhee
Alex Pitt
Lance Munday

7.6 Monitoring and evaluation

Project management was as planned at the start of the program, with six-monthly written reports to Wine Australia and steering committee meetings with Wine Australia, the industry reference group and the leaders of the other projects within the RRD4P program led by Wine Australia every six to twelve months. These were normally hosted by CSIRO Agriculture and Food and included detailed presentations on work since the previous steering group meeting. The meetings were useful to get some industry feedback on the project and this was used to help focus the direction of the work.

The project has successfully demonstrated a range of technology applications as detailed above and this has helped the program achieve its objectives of generating new knowledge and technologies to benefit primary producers and to build collaborative linkages between research and industry.

7.7 Budget

A budget reconciliation will be provided separately.