# Highlights

**The influence of resource use on yield versus quality trade-off in Australian vineyards**

Author

- Comparative analysis of resource use, quality and quantity in Australian winegrowing.
- Regional comparison of outcomes and resource use in Australian winegrowing regions.
- Baseline models for comparing wine crops.
- Analysis of national, decade long data source.

# The influence of resource use on yield versus quality trade-off in Australian vineyards

Author[1,1,1]

## Abstract

When strategies for a sustainable winegrowing industry are assessed, there is a trade-off between balancing the amount of resources invested and the resultant yield and quality of the produce. In this analysis we observe relationships between resource use, yield and quality through the use of statistical models. The dataset used for this analysis includes data collected for the past 10 years from a multitude of vineyards located over a diverse range of Australian winegrowing regions. Yield and quality (measured as a ratio of sale price to area) was modelled to resource factors related to water usage and emissions. The analysis confirmed an expected strong relationship between size and resource use, with the overall space of a vineyard and its access to resources greatly determining the upper limit of potential yield. However, size was also negatively related to the potential quality, with higher quality being connected to high resource inputs per area; rather than to the overall expenditure of resources. Regional and yearly effects on Vineyard outputs were also identified. Overall, the analysis highlighted the importance of considering a vineyard's business goal, region, external pressures and economies of scale, with regional constraints also contributing to deciding the best resource use strategies to pursue when considering quality or quantity.

## 1. Introduction

The global focus on sustainability in agronomic industries has changed the way in which these enterprises do business. When strategies for a sustainable winegrowing industry are assessed, there is a trade-off between balancing the amount of resources invested and the resultant yield versus quality produced. This dilemma exists across agriculture through shared fundamental considerations such as water use and fuel usage (Hemming et al., 2020; Kawasaki and Uchida, 2016; ZHU et al., 2017). Quality in viticulture (the cultivation of grapes for wine production) is driven through its integration within the wine industry, with the potential quality of a wine being initially defined through the chemical makeup of the grapes used in its production. The consideration of sustainability within viticulture is further complicated by environmental and socio-demographic pressures. In the Australian context, these include biosecurity, climate and international market demands.

There is an extensive amount of research into the effects of a variety of factors on grape quality and yield (He et al., 2022; Laurent et al., 2022; Liakos et al., 2018). However, due to the lack of long-term and in-depth data, individual factors are often studied in isolation (Abbal et al., 2016). The lack of consolidated datasets also restricts the ability to gain statistical insights at large scales and across multiple regions (Keith Jones, 2002; Knight et al., 2019). The dataset used for this analysis includes data collected for the past 10 years from a multitude of vineyards located over a diverse range of Australian winegrowing regions. We aim to use this dataset to describe the relationship of resources related to water and fuel use with the output yield and quality of the resultant product, taking into account

Table 1: Summary of models; their predictors, covariates and variable interactions.

| | Response | Predictors | Covariates | Interactions |
|---|---|---|---|---|
| **Model 1** | Yield | Water Used<br>scope one Emissions | Area Harvested<br>Year<br>GI Region | N/A |
| **Model 2** | $\frac{\text{Yield}}{\text{Area Harvested}}$ | Water Used<br>scope one Emissions | Area Harvested<br>Year<br>GI Region | Area Harvested * scope one Emissions<br>Area Harvested * Water Use<br>Year * Region |
| **Model 3** | Yield×Average Sale Price | Water Used<br>Scope One Emissions | Area Harvested<br>Year<br>GI Region | N/A |
| **Model 4** | Average Sale Price | Water Used<br>Scope One Emissions | Area Harvested<br>Year<br>GI Region | Area Harvested * Scope One Emissions<br>Area Harvested * Water Use<br>Year * Region |
| **Model 5** | Average Sale Price | Water Used<br>Scope One Emissions | Year<br>GI Region | Year * Region |

the size and location of the vineyard. The practical addition of this aim is a baseline for comparison: given a vineyard within Australia, one could estimate the comparative efficiency with regard to the tradeoff between invested resources, yield and quality. This is the first time that such a trade off has been confirmed explicitly across such varying regions, scales and climates in the Australian winegrowing industry.

## 2. Methods

### *2.1. Data*

Data used in this analysis were obtained from Sustainable Winegrowing Australia and Wine Australia. Sustainable Winegrowing Australia is

Australia's national wine industry sustainability program, which aims to facilitate grape-growers and winemakers in demonstrating and improving their sustainability (SWA, 2022). Wine Australia is an Australian Government statutory authority governed by the Wine Australia Act 2013 (Win, 2019).

Predictor variables used in this analysis included yield, defined as the total tonnes of grapes harvested, and quality, defined as average sale price of grapes. It is acknowledged that quality can be defined in a variety of ways, for example by the grapes': aroma, chemical composition and color (Kasimati et al., 2022; Mejean Perrot et al., 2022; Suarez et al., 2021). Using sale price was based on the reliance of market value of winegrapes on grape quality and because Wine Australia explicitly defines grape quality through the use of discrete price brackets in their annual reports. The generalisation made to reflect quality through using average price assumed a due diligence of those who purchased the grapes (Yegge, 2001). Both response variables were examined as totals and as scales of area harvested. Values were compared in this manner to observe how economies of scale affect the use of resources.

Data obtained from Wine Australia were collected via phone surveys and included: total tonnes purchased, average price per tonne and yearly change in price for region and grape varietal. Data recorded by Sustainable Winegrowing Australia was entered manually by winegrowers using a web based interface with some fields being optional. Required variables included: region, harvest year, yield and area harvested. Optional variables included average sale price, water used and fuel used (diesel, petrol, biodiesel and LPG). To enable direct comparisons between fuels, fuel use was converted to tonnes of Carbon Dioxide equivalent and collectively referenced to as emissions.

Average sale price was an optional field in the Sustainable Winegrowing Australia's dataset. Missing values were improved using regional average prices from Wine Australia. Two subsets of data were then created for the analysis. The first subset contained all vineyards and was used for two models (Model 1 and Model 2, see Table 1). The second subset contained vineyards which either recorded a value for average price of sale per tonne through Sustainable Winegrowing Australia, or were within a region with an average price of sale recorded by Wine Australia; this subset was used for three further models (Models 3, 4 and 5, see Table 1. These subsets meant that the data would be limited to samples which had recorded values for the response variables (see Table 1), where every sample had a recorded value for yield but not average price of sale per tonne.

The first subset of data (used for Model 1 and Model 2, see Table 1) contained 5298 samples spanning the period from 2012 to 2022, covering 55 GI Regions and 1261 separate vineyards.

The second subset of data (used for Model 3, Model 4 and Model 5, see Table 1) contained 2878 samples spanning the period from 2015 to 2022, covering 51 GI Regions and 944 separate vineyards.Average price of sale per tonne was extracted from both Wine Australia (1842 values) and Sustainable Winegrowing Australia (remaining 1036 values).

Additional variables were considered for analysis but were excluded due to being either underreported or had insignificant contributions to model accuracies. Variables explored but not used due to low reporting values included fertiliser, and scope two emissions. Variables considered but ultimately removed due to a lack of significant contributions to models, included the use of renewable energy, contractor use, and pressures such as frost, fire and disease.

Data preprocessing was conducted prior to analysis using the Python programming language (G. van Rossum, 1995). Preprocessing included the conversion from fuel to scope one emissions and prior calculations for all continuous variables which included logarithmic transformations, centring and scaling by standard deviation. The transformation of fuel use into scope one emissions was done using the equation given from the Australian National Greenhouse Accounts Factors, shown as

$$tCO_2e = \frac{Q \times EC \times EF1 + EF3}{1000}, \tag{1}$$

was used to convert the quantity of fuel in litres, $Q$, using a prescribed Energy Content, $EC$, and emission factors of scope one, $EF1$, and scope three, $EF3$, to tonnes of Carbon Dioxide Emission equivalent, $tCO2e$ (Department of Climate Change, Energy, the Environment and Water, 2022).

Differences in vineyard locations were captured through the use of Geographical Indicator Regions (GI Regions, see Figure 1). Each GI Region has its own unique mixture of climatic and geophysical properties that describes a unique winegrowing region within Australia; these regions were predefined by Wine Australia (Halliday, 2009; Oliver et al., 2013; SOAR et al., 2008). Both Wine Australia and Sustainable Winegrowing Australia used the same GI Region format to describe location.

The site of a vineyard predetermines several physical parameters such as climate, geology and soil, making location a widely considered key determi-

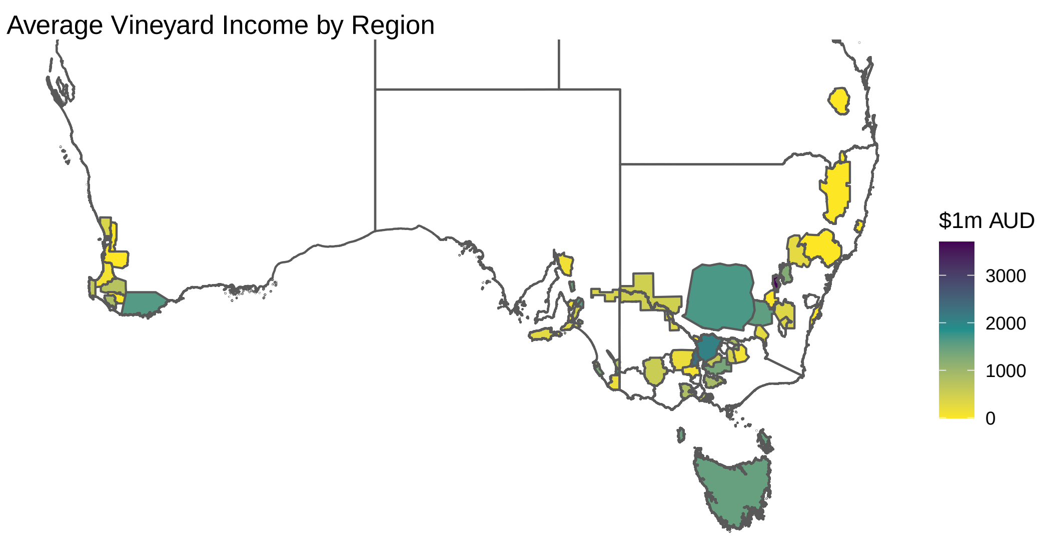


Figure 1: Map of vineyard average income for each of the used GI Regions.

nant of grape yield and quality (Abbal et al., 2016; Agosta et al., 2012; Fraga et al., 2017). The climatic properties of each GI Region were summarised by using predefined classifications as per the Sustainable Winegrowing Australia (2021) user manual. The user manual describes climates by rainfall and temperature, creating supersets of Regions of similar climatic properties. The climatic groups were used to illustrate similarities and differences occurring in areas larger than GI Regions.

### *2.2. Analysis*

Pairwise Pearson Correlation Coefficients were calculated to assess the potential existence of linear relationships between the input and predicted variables. To determine if a coefficient was indicative of a strong relationship, confidence intervals were used. P-values reflected the significance of a given correlation coefficient with statistical significance being declared when the associated value was lower than 0.05. Pairwise Pearson Correlation Coefficients were calculated for data on the original scale and for data as a logarithmic transform. Transforming data prior to calculating the coefficients changes several things. The logarithmic transform of the data alters the interpretation of the coefficients to percentage change; a coefficient will be indicative of the change in percentage of one variable compared to the other, scaling by standard deviation also changes this interpretation to be a percentage of that variables standard deviation. When considering the logarithmically transformed variables, a coefficient of 1 would indicate that the change of one variable by one percentage of its standard deviation would correlate to the other variable changing by one percent of its own standard deviation. The importance of this is the dimensionless nature of these relationships and that it can be translated directly to any vineyard's case that has a well known distribution.

Five general linear models were created (see Table 1). General Linear Models were chosen as they offer the ability to produce statistical models that are explicit in the relationships between predictors and response variables. General Linear Models also allowed the exploration of interactions between predictors and allow for easily comparable differences in the influence and magnitude of relationships. Model fit was measured in $R^2$ and adjusted $R^2$ as well as F statistics. T-tests were used to determine if predictors significantly contributed to their models when accounting for other variables, showing which specific years and areas contributed significantly. Both the Pearson Correlation Coefficients and General Linear Models were created using the R statistical programming language (R Core Team, 2021) with the Caret package (Kuhn, 2008).

A variety of alternate methods were also explored, including splines, hierarchical regression, General Additive Models, and Generalised Linear Models. These alternative approaches were not used as final models due to offering no further insights or improvements in accuracy.

### *2.3. Model Validation*

Models were validated using K-fold cross validation calculated. K-fold cross validation works by removing a subset of data from the sample used to train models and then predicts those variables to determine how sensitive the model is to changes in the sample data. For this analysis each model was validated using 10 folds, repeated 100 times.

Table 2: Summary statistics of each continuous variable.

| **Variable** | **Mean** | **Standard Deviation** | **Minimum** | **Maximum** |
| --- | --- | --- | --- | --- |
| Yield (tonnes) | 7.757E+02 | 2.179E+03 | 1.000E+00 | 7.231E+04 |
| Area Harvested (ha) | 6.670E+01 | 1.337E+02 | 7.000E-02 | 2.436E+03 |
| Water Used (ML) | 7.471E+06 | 5.646E+08 | 1.000E+00 | 4.268E+10 |
| Scope One Emissions ($tCO_2e$) | 4.173E+04 | 8.571E+04 | 6.755E+00 | 2.110E+06 |
| $\frac{\text{Yield (tonnes)}}{\text{Area harvested (ha)}}$ | 1.009E+01 | 8.127E+00 | 4.000E-02 | 8.634E+01 |
| Average Sale Price (AUD) | 1.477E+03 | 9.216E+02 | 1.600E+02 | 2.600E+04 |
| $\frac{\text{Average Sale Price (AUD)}}{\text{Area Harvested (ha)}}$ | 1.347E+02 | 5.711E+02 | 1.753E-01 | 2.979E+04 |

# 3. Results

## *3.1. Exploratory Analysis*

Table 2 shows the summary statistics of each variable in its original units. The range of these values shows the level of difference between some vineyards, with operations differing by orders of magnitude in size, yield and average price of sale (See Table 1).

Pearson Correlation Coefficients of the transformed, centred and scaled

Table 3: Pairwise Pearson correlation coefficients for logarithmically transformed values.

| | **Yield** | **Area Harvested** | **Water Used** | **Scope One Emissions** | **Yield by Area** | **Average Price** | **Average Price by Area** |
|---|---|---|---|---|---|---|---|
| Yield | 1.00 | 0.88 | 0.82 | 0.76 | 0.96 | -0.46 | -0.88 |
| Area Harvested | 0.88 | 1.00 | 0.78 | 0.83 | 0.73 | -0.19 | -0.81 |
| Water Used | 0.82 | 0.78 | 1.00 | 0.67 | 0.76 | -0.49 | -0.82 |
| Scope One Emissions | 0.76 | 0.83 | 0.67 | 1.00 | 0.65 | -0.16 | -0.67 |
| Yield by Area | 0.96 | 0.73 | 0.76 | 0.65 | 1.00 | -0.54 | -0.84 |
| Average Price | -0.46 | -0.19 | -0.49 | -0.16 | -0.54 | 1.00 | 0.72 |
| Average Price by Area | -0.88 | -0.81 | -0.82 | -0.67 | -0.84 | 0.72 | 1.00 |

variables are shown in Table 3. All correlations were found to be statistically significant (P $<$ 2.200E-16), and except for 'average price' all variables were positively correlated. With water use, area harvested and emissions being positively correlated to yield, it can be considered that more resources and area are likely to lead to greater yields. Average sale price's negative correlation to yield, water use, area and scope one emissions, indicated that size and fuel separately were not the determining factor for grape quality. The negative correlations are not causal relationships (using more water does not cause lower quality) but relative are measures indicating that using greater amounts of water than others may lead to lower quality.

### *3.2. General Linear Models*

Each model had a high $R^2$ value, indicating that a most of the variance within the data was described by the models (see Table 4). The models were found to be a good fit, with overall F-tests being statistically significant (P $<$ 2.200E-16). And, aside from 3 variables, F-tests across each model's variables

Table 4: Summary of models; their performance, F-statistics and Residual error.

| | $R^2$ | Adjusted $R^2$ | F-Statistic | P-Value | Residual Standard Error | Residual Sum of Squares | Residual Mean of Squares |
|---|---|---|---|---|---|---|---|
| **Model 1** | 0.9072 | 0.9061 | 775.3 | 2.200e-16 | 0.3065 | 491.3 | 0.1 |
| **Model 2** | 0.8291 | 0.8141 | 55.07 | 2.200e-16 | 0.4312 | 905.03 | 0.19 |
| **Model 3** | 0.9753 | 0.9748 | 1885 | 2.200e-16 | 0.1589 | 71.11 | 0.03 |
| **Model 4** | 0.9091 | 0.9006 | 106.1 | 2.200e-16 | 0.3153 | 261.41 | 0.10 |
| **Model 5** | 0.9089 | 0.9004 | 107.2 | 2.200e-16 | 0.3155 | 262.04 | 0.10 |

were also significant (with all being at least, $P < 0.05$). The three exceptions were: scope one emissions in Model 3 (P=0.22) and Model 4 (P=0.0.39), and the interaction between area harvested and water used in model 2 (P=0.22). Note that, scope one emissions was included in all models to directly compare the response variables as ratios of vineyard size to raw values and because it was strongly correlated to the response variable in every model (except model 5); especially for Models 1 and 4 (Table 3). A full list of regression coefficients 95% CIs and p-values for each of the four models is provided in the appendix.

Models' continuous variable's coefficient values are summarised in Table 5. Model 1 showed all coefficients except for the intercept were significantly contributing to the model (P ¡ 0.05). Model 2's coefficients were all statistically significant. However, for Models 3, 4 and 5 Scope one emissions did not significantly contribute. And, Model 4 only saw statistically significant contributions from the intercept and water use. Although the coefficient for water use was statistically significant for each model, it did not have the

Table 5: Summary of each Models coefficients for continuous variables

| | | Intercept | Area Harvested | Water Used | Scope One Emissions | Area Harvested Scope One Emissions | Area Harvested Water Used |
|---|---|---|---|---|---|---|---|
| Model 1 | Coefficient | -0.0332 | 0.7418 | 0.0866 | 0.0673 | | |
| | Std Error | 0.0196 | 0.0100 | 0.0089 | 0.0080 | | |
| Model 2 | Coefficient | 0.1696 | 0.5774 | 0.1079 | 0.0850 | -0.0497 | -0.0535 |
| | Std Error | 0.0591 | 0.0148 | 0.0131 | 0.0117 | 0.0081 | 0.0084 |
| Model 3 | Coefficient | 0.0181 | 0.9713 | -0.0231 | -0.0070 | | |
| | Std Error | 0.0130 | 0.0072 | 0.0069 | 0.0057 | | |
| Model 4 | Coefficient | 0.1450 | 0.0024 | -0.0466 | -0.0170 | 0.0115 | 0.0014 |
| | Std Error | 0.0528 | 0.0150 | 0.0143 | 0.0118 | 0.0079 | 0.0083 |
| Model 5 | Coefficient | 0.1517 | | -0.0404 | -0.0171 | | |
| | Std Error | 0.0527 | | 0.0113 | 0.0097 | | |

highest value, instead area harvested, being an order of magnitude greater dominated the models. Model 5 was able to achieve a similar $R^2$ to Model 4 without area harvested, having stronger influences from water use and scope one emissions.

The regression coefficients for the year for each model is depicted in Figure 2. The first year for a model's data is used as the baseline. The Adelaide Hills is used as the regional baseline with the interaction between year and region using the first year and the Adelaide Hills as the baseline. Region and year contributed, in some but not all cases, more than the other variables. However, some years are not significant, as they are not statistically different from 0, given their error. Models 4 and 5 are very similar, indicating that the exclusion of area does not greatly affect the contribution from yearly influence. Models 4 and 5 have the most prominent trends, showing an increase in yearly effects over time, with Model 3 also increasing from 2016 to 2018 but plateau afterwards. Models 1 and 2 do not show a clear trend but do drop during 2017 and 2018 after increasing in the first 3 years.

Regional differences are summarised in Figure 3. The most notable difference is between vineyards within 'Hot' and 'Very Dry' regions (warm inland regions), where little emphasis is put on achieving high average sale prices, instead focussing on larger scale yield. Water Use changes dramatically between these regions as well, with water being a driving force in the mass production of grapes but not necessarily the quality. The warmer and drier regions tend to also cater to larger vineyards, with greater areas.

Figure 4 further shows the emphasis that 'Hot' areas have on high yields with low average sale price compared with other regions. Scaling average price and yield by area shows a strong negative trend, trading quantity for higher sales prices.

Table 3.2 shows the validation results of each of the models. The $R^2$ measures of fit show similar results to the initial models, with a slight decrease. Indicating that the models are robust and consistent.

## 4. Discussion

There was an expected strong relationship between size and resource use, with the overall space of a vineyard and its access to resources greatly determining the upper limit of potential yield. However, size was also inversely related to the potential quality, with higher quality being related to high resource inputs per area; rather than to the overall expenditure of resources. Vineyard outputs were also augmented by regional and yearly affects. Even

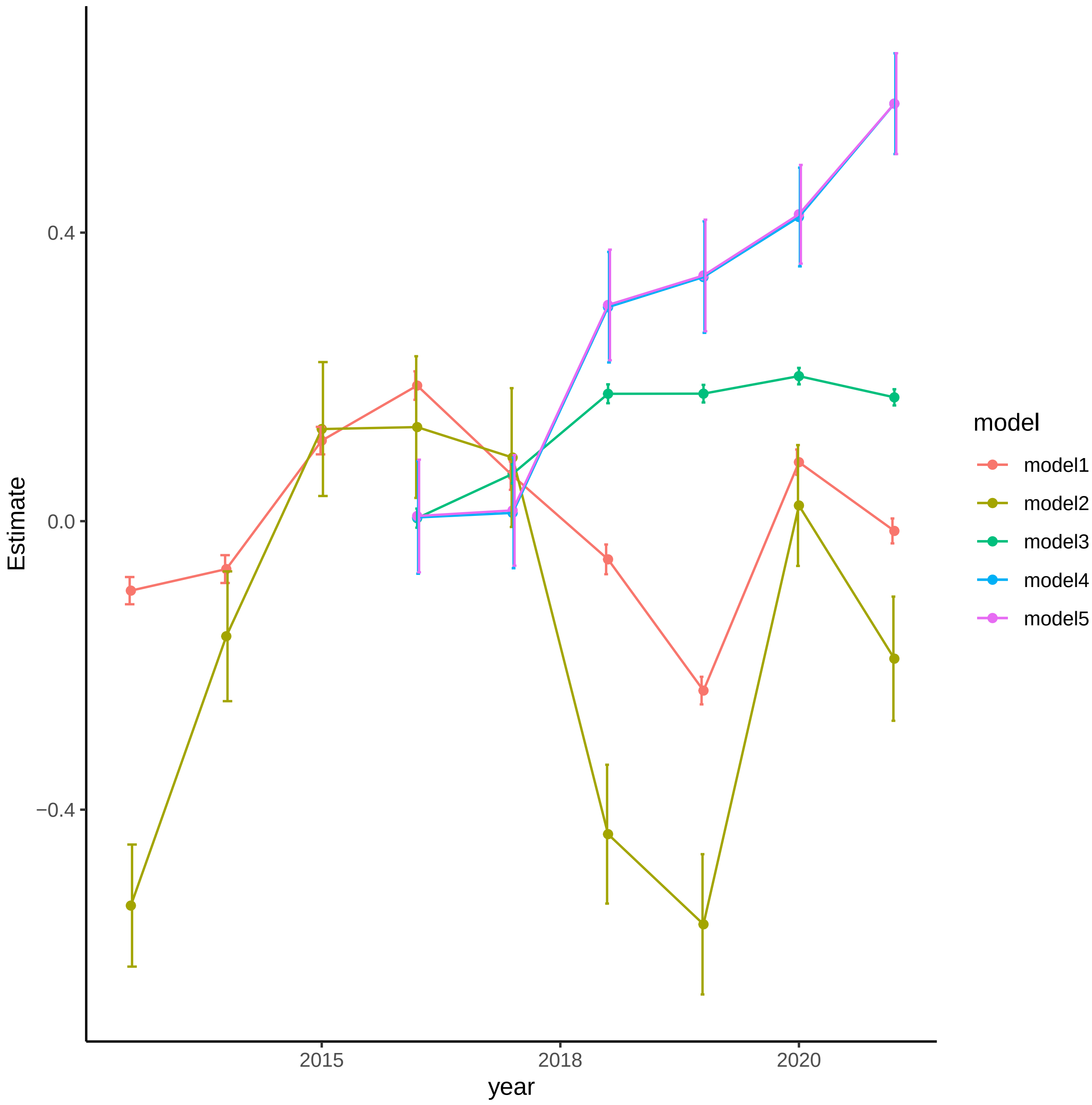


Figure 2: Model Coefficient values for Year, with standard error bars.

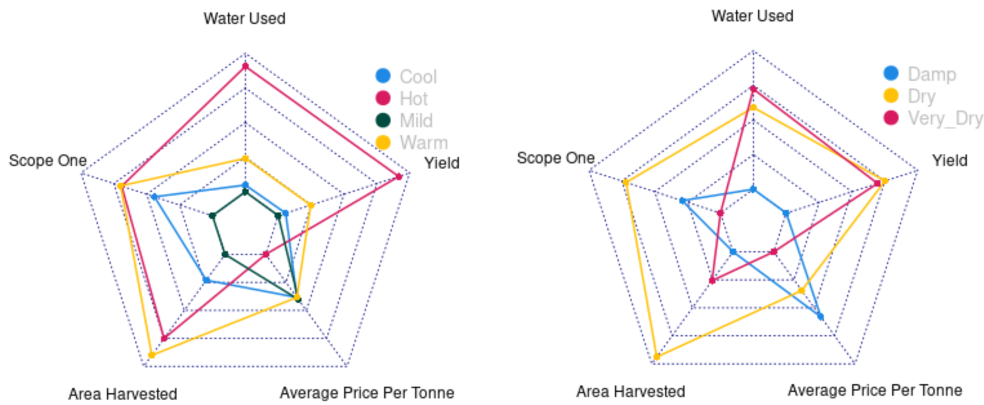


Figure 3: Radar plot of climatic profile's resource use, yield and average sale price.

Table 6: Model validation using k-fold cross validation, for 10 folds repeated 100 times.

| | **Residual Mean Squared Error** | **R2** | **Mean Average Error** |
|---|---|---|---|
| **Model 1** | .309 | .905 | .2165 |
| **Model 2** | .457 | .7921 | .313 |
| **Model 3** | .165 | .972 | .101 |
| **Model 4** | .348 | .878 | .182 |
| **Model 5** | .348 | .878 | .183 |

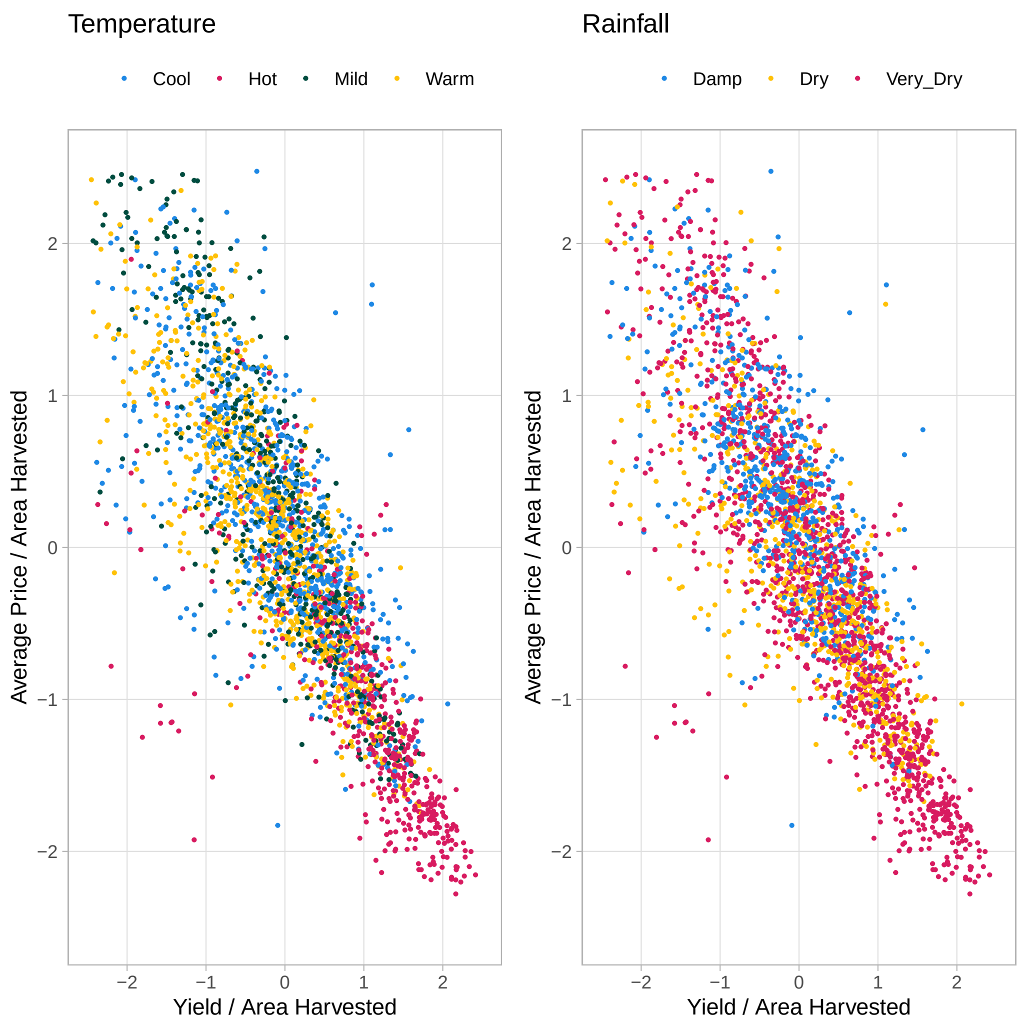


Figure 4: Scatter plot of vineyard yield against the average sale price as ratios to area harvested. The axes are in standard deviations with points coloured by climate.

given regional and yearly changes, there was a strong connection between smaller vineyards and higher quality. This could have been due to the easier management of smaller properties.

Scope one emissions' lack of significance and contribution given its F-statistics, could be indicative that other vineyard activities requiring fuel are not leading factors for a vineyards grape quality. The relationship between yield, value and area was not simply about efficiently producing the most grapes. It is possible that the relationship of scope one emissions between yield and sale price was closely tied to a vineyard's area due to requiring more fuel to address more issues over greater distances. It is difficult to discern the connection of scope one emissions directly, as fuel can be used for a broad category of activities.

There are important considerations unique to winegrowing compared to other agricultural industries. The vertical integration of winegrowing within the wine industry ties winegrowers to secondary and tertiary industries, such as wine production, packaging, transport and sales. This results in unique issues and considerations for each vineyard, where on-the-ground decisions are influenced by other wine industry's choices, such as the use of sustainable practices in vineyards as a requirement for sale in overseas markets; notably these interactions can be further complicated by some winegrowers being completely integrated into a wine company, while others are not (Knight et al., 2019). Incorporating decisions into the model could help describe the contributing factors to regional differences beyond resource consumption, motivating the call for more granular data and more sophisticated modelling.

There are many on-the-ground decisions that influence both quality and yield. The decision to prioritise quality over quantity, is governed by complex physical and social forces, for example international market demands, disease pressures and natural disasters (Abad et al., 2021; Cortez et al., 2009; Hall et al., 2011; I. Goodwin, et al., 2009; Kasimati et al., 2022; Oliver et al., 2013; Srivastava and Sadistap, 2018), with many of these occurrences being highlighted throughout the reports from Wine Australia (Wine Australia, 2019, 2021, 2022; Winemakers' Federation of Australia, 2013, 2014, 2015, 2016, 2017, 2018) over the past decade. However, the changes in the coefficients (see Figure 2) are not reflective of many known occurrences, such as the 2020 bush fires, which had higher values for coefficients than prior years; During the 2020 bush fires 40,000 tonnes of grapes were lost across 18 different wine regions due to bush fires and smoke taint. In comparison to countrywide pressures such as drought, this damage made up only 3% of the total amount of grapes for that year; although acknowledged as a considerable loss on an individual basis, it was deemed to be only a minor national concern by Wine Australia when compared to other environmental pressures such as drought (Wine Australia, 2020)

Climatic pressures are an important consideration for growers, especially those in warmer and drier regions. The Wine Australia reports also show that warm inland regions have seen a decline in profit over the past decade, whereas regions targeting quality did not. The warm inland regions also tend to contain larger vineyards, making up for lower sale prices with larger yields. Considering the negative correlation of average price to area, for this strategy to work economies of scale become and important factor. Given the large quantities of grapes that can be produced by some vineyards, even at low margins there is the potential to be profitable. However, the increasing climatic pressures mixed with the requirement for larger volumes of water, make the sustainability of some vineyards come into question. Furthermore, intensive farming in general is known to jeopardise the sustainability of an operation through the degradation of soil and waterways (Capello et al., 2019; Lin, 2012; Pisciotta et al., 2015). There are established methods that can help to mitigate these affects, such as the use of cover crops and crop rotation. However, it has become more apparent that the active reduction of grape yield, through methods such as thinning, can help increase the quality of grapes and improve soil health (Condurso et al., 2016; Wang et al., 2019).

Some regions appeared to produce many low quality grapes at scale whilst others focussed on producing higher quality grapes in lower volumes. This emperical finding is consistent with Wine Australia's annual reports, which shows that some GI regions, such as the Riverland, are known for producing large amounts of lower grade (low value per tonne) grapes (Wine Australia, 2022; Winemakers' Federation of Australia, 2017). Comparatively, other regions, such as Tasmania, only produce high quality grapes but in smaller quantities. The difference in pricing per tonne between the lowest and highest graded grapes can be greater than a hundred times the difference in value per tonne. Not all regions target only one grade of grape, with some producing a variety of differently graded grapes; such as the Yarra Valley, which produces grades from C to A. This effect could stifle the potential for market opportunities within lesser known regions. A further possibility is the existence of regional upper limits on potential quality, or that there are diminishing returns in some regions when pursuing quality or quantity;

however these types of relationships may be obfuscated by knowledgeable winegrowers who avoid this pitfall.

Due to regional differences, different strategies are also employed across different regions, such as some regions targeting mass production over quality. This is most notable when grouping regions by climate, especially when considering GI Regions in the 'Hot Very Dry' climate (see Figure 4). Although not chosen over GI region, climate was considered to be a large determinant of the ability to produce larger quantities of grapes, as well as a determinant in grape quality (Agosta et al., 2012). The more granular GI Region likely explained a broader mix of geographical phenomenon, such as soil, geology and access to water resources (Abbal et al., 2016; Carmona et al., 2011). The interaction between year and GI Region likely accounted for events such as bushfires, which would be impactful, but only at a local level, both in time and space.

Limitations in the analyses presented in this paper included overestimating yield for models 1 and 2, and underestimating crop value in models 3 and 4 (see appendix). The issue of model 1 and 2 over-predicting yield may have been due to preventative measures brought on by regional pressures such as fire, frost and disease. More resources were required to prevent these issues from spreading within a region, thus disproportionately affecting some vineyards compared to others locally. This type of maintenance is not well captured in the models, especially when considering that some regions, especially those in warmer areas, are not as prone to disease as cooler climates and could potentially have lower operating costs per hectare. This could create a discrepancy in vineyards that utilised preventative measures in wetter regions, as opposed to those that did not, thus expending less fuel and energy but risking disease. When reviewing the differences between regions, it is important to consider that vineyards in 'Hot Very Dry' areas can be hundreds of times the size of those in other regions. This limitation could be overcome by incorporating the profitability of vineyards, comparing the financial success of working at different operational scales.

Variables such as the utilisation of renewable energy, contractors, and the occurrence of disease, fire and frost were originally explored to capture the discrepancies between similar vineyards that produced different yields and crop values. However, none of these variables was significantly correlated with the response variables, and did not add to model accuracy, even when considered as interactions. Allowance for nonlinear relationships, specifically through splines, resulted in more normally distributed residuals but at a drastically reduced overall accuracy when comparing $R^2$ and Residual Square Error. Attempts to fully explain small variations was always overshadowed by the dramatic differences in regional trends.

Having more data for each region would also be beneficial, allowing greater comparison between regions. More variables may also help to discern vineyards that can produce larger volumes of grapes at higher prices. The use of other models such as random forests and decision trees alongside more variables and data may help to uncover the reasons for under or overestimation. These differences could be caused by the use of alternative sustainable practices in the field. Moreover, while there is evidence to suggest that environmentally sustainable practices can reduce costs, and increase efficiency whilst improving the quality of grapes; more research is needed to link these benefits across different regions and climates (Baiano, 2021; Mariani and Vastola, 2015; Montalvo-Falcón et al., 2023).

## 5. Conclusion

This study delved into the relationships between resource use, grape quality and yield. The findings underscore the multifaceted nature of vineyard management, where the interplay of size, resource allocation, climate, and regional influences collectively shape both the quality and quantity of grape yields. Quality was not solely tied to the overall expenditure of resources, but rather to the efficient allocation of resources per unit area. This emphasises that factors beyond sheer scale contribute significantly to the final quality of the grapes produced. Moreover, regional and yearly variations exhibited substantial effects on vineyard outputs, impacting both quantity and quality. The connection observed between smaller vineyards and higher grape quality suggests that the management of smaller properties might be more streamlined and effective, enabling a greater quality of grape to be produced.

## References

, 2019. Wine Australia Act 2013.

Abad, J., Hermoso de Mendoza, I., Marín, D., Orcaray, L., Santesteban, L.G., 2021. Cover crops in viticulture. A systematic review (1): <br>Implications on soil characteristics and biodiversity in vineyard. OENO One 55, 295–312. doi:10.20870/oeno-one.2021.55.1.3599.

Abbal, P., Sablayrolles, J.M., Matzner-Lober, É., Boursiquot, J.M., Baudrit, C., Carbonneau, A., 2016. Decision Support System for Vine Growers

Based on a Bayesian Network. Journal of agricultural, biological, and environmental statistics 21, 131–151. doi:10.1007/s13253-015-0233-2.

Agosta, E., Canziani, P., Cavagnaro, M., 2012. Regional climate variability impacts on the annual grape yield in Mendoza, Argentina. Journal of Applied Meteorology and Climatology 51, 993–1009.

Baiano, A., 2021. An Overview on Sustainability in the Wine Production Chain. Beverages 7. doi:10.3390/beverages7010015.

Capello, G., Biddoccu, M., Ferraris, S., Cavallo, E., 2019. Effects of Tractor Passes on Hydrological and Soil Erosion Processes in Tilled and Grassed Vineyards. Water 11. doi:10.3390/w11102118.

Carmona, G., Varela-Ortega, C., Bromley, J., 2011. The Use of Participatory Object-Oriented Bayesian Networks and Agro-Economic Models for Groundwater Management in Spain. Water resources management 25, 1509–1524. doi:10.1007/s11269-010-9757-y.

Condurso, C., Cincotta, F., Tripodi, G., Sparacio, A., Giglio, D.M.L., Sparla, S., Verzera, A., 2016. Effects of cluster thinning on wine quality of Syrah cultivar (Vitis vinifera L.). European food research & technology 242, 1719–1726. doi:10.1007/s00217-016-2671-7.

Cortez, P., Teixeira, J., Cerdeira, A., Almeida, F., Matos, T., Reis, J., 2009. Using data mining for wine quality assessment, in: Discovery Science: 12th International Conference, DS 2009, Porto, Portugal, October 3-5, 2009 12, Springer. pp. 66–79.

Department of Climate Change, Energy, the Environment and Water, 2022. Australian National Greenhouse Accounts Factors.

Fraga, H., Costa, R., Santos, J.A., 2017. Multivariate clustering of viticultural terroirs in the Douro winemaking region. Ciência Téc. Vitiv. 32, 142–153.

G. van Rossum, 1995. Python tutorial, Technical Report CS-R9526. Centrum voor Wiskunde en Informatica (CWI),.

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

Halliday, J.C.J.C., 2009. Australian Wine Encyclopedia. Hardie Grant Books, VIC.

He, L., Fang, W., Zhao, G., Wu, Z., Fu, L., Li, R., Majeed, Y., Dhupia, J., 2022. Fruit yield prediction and estimation in orchards: A state-of-the-art comprehensive review for both direct and indirect methods. Computers and Electronics in Agriculture 195, 106812. doi:10.1016/j.compag.2022.106812.

Hemming, S., de Zwart, F., Elings, A., Petropoulou, A., Righini, I., 2020. Cherry tomato production in intelligent greenhouses-sensors and ai for control of climate, irrigation, crop yield, and quality. Sensors (Basel, Switzerland) 20, 1–30. doi:10.3390/s20226430.

I. Goodwin,, L. McClymont,, D. Lanyon, A. Zerihun, J. Hornbuckle, M.

Gibberd, D. Mowat, D. Smith, M. Barnes, R. Correll, 2009. Managing soil and water to target quality and reduce environmental impact.

Kasimati, A., Espejo-García, B., Darra, N., Fountas, S., 2022. Predicting Grape Sugar Content under Quality Attributes Using Normalized Difference Vegetation Index Data and Automated Machine Learning. Sensors 22. doi:10.3390/s22093249.

Kawasaki, K., Uchida, S., 2016. Quality Matters More Than Quantity: Asymmetric Temperature Effects on Crop Yield and Quality Grade. American journal of agricultural economics 98, 1195–1209. doi:10.1093/ajae/aaw036.

Keith Jones, 2002. Australian Wine Industry Environment Strategy.

Knight, H., Megicks, P., Agarwal, S., Leenders, M., 2019. Firm resources and the development of environmental sustainability among small and medium-sized enterprises: Evidence from the Australian wine industry. Business Strategy and the Environment 28, 25–39. doi:10.1002/bse.2178.

Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. Journal of Statistical Software, Articles 28, 1–26. doi:10.18637/jss.v028.i05.

Laurent, C., Le Moguédec, G., Taylor, J., Scholasch, T., Tisseyre, B., Metay, A., 2022. Local influence of climate on grapevine: An analytical process involving a functional and Bayesian exploration of farm data time series synchronised with an eGDD thermal index. OENO one 56, 301–317. doi:10.20870/oeno-one.2022.56.2.5443.

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine Learning in Agriculture: A Review. Sensors 18. doi:10.3390/s18082674.

Lin, H., 2012. Hydropedology : Synergistic Integration of Soil Science and Hydrology. Elsevier Science & Technology, San Diego, NETHERLANDS, THE.

Mariani, A., Vastola, A., 2015. Sustainable winegrowing: Current perspectives. International Journal of Wine Research 7, 37–48.

Mejean Perrot, N., Tonda, A., Brunetti, I., Guillemin, H., Perret, B., Goulet, E., Guerin, L., Picque, D., 2022. A decision-support system to predict grape berry quality and wine potential for a Chenin vineyard. Computers and electronics in agriculture 200, 107167. doi:10.1016/j.compag.2022.107167.

Montalvo-Falcón, J.V., Sánchez-García, E., Marco-Lajara, B., Martínez-Falcó, J., 2023. Sustainability Research in the Wine Industry: A Bibliometric Approach. Agronomy 13. doi:10.3390/agronomy13030871.

Oliver, D., Bramley, R., Riches, D., Porter, I., Edwards, J., 2013. Review: Soil physical and chemical properties as indicators of soil quality in Australian viticulture. Australian Journal of Grape and Wine Research 19, 129–139. doi:10.1111/ajgw.12016.

Pisciotta, A., Cusimano, G., Favara, R., 2015. Groundwater nitrate risk assessment using intrinsic vulnerability methods: A comparative study of environmental impact by intensive farming in the Mediterranean region of Sicily, Italy. Journal of geochemical exploration 156, 89–100. doi:10.1016/j.gexplo.2015.05.002.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

SOAR, C., SADRAS, V., PETRIE, P., 2008. Climate drivers of red wine quality in four contrasting Australian wine regions. Australian journal of grape and wine research 14, 78–90. doi:10.1111/j.1755-0238.2008.00011.x.

Srivastava, S., Sadistap, S., 2018. Non-destructive sensing methods for quality assessment of on-tree fruits: A review. Journal of Food Measurement and Characterization 12, 497–526.

Suarez, L., Zhang, P., Sun, J., Wang, Y., Poblete, T., Hornero, A., Zarco-Tejada, P., 2021. Assessing wine grape quality parameters using plant traits derived from physical model inversion of hyperspectral imagery. Agricultural and forest meteorology 306, 108445. doi:10.1016/j.agrformet.2021.108445.

Sustainable Winegrowing Australia, SWA., 2021. Sustainable Winegrowing Australia User Manual.

SWA, S.W.A., 2022. Sustainable Wingrowing Australia. https://sustainablewinegrowing.com.au/case-studies/.

Wang, Y., He, Y.N., He, L., He, F., Chen, W., Duan, C.Q., Wang, J., 2019. Changes in global aroma profiles of Cabernet Sauvignon in response to cluster thinning. Food research international 122, 56–65. doi:10.1016/j.foodres.2019.03.061.

Wine Australia, 2019. National Vintage Report 2019 .

Wine Australia, 2020. National Vintage Report 2020 .

Wine Australia, 2021. National Vintage Report 2021 .

Wine Australia, 2022. National Vintage Report 2022 .

Winemakers' Federation of Australia, 2013. National Vintage Report 2013 .

Winemakers' Federation of Australia, 2014. National Vintage Report 2014 .

Winemakers' Federation of Australia, 2015. National Vintage Report 2015 .

Winemakers' Federation of Australia, 2016. National Vintage Report 2016 .

Winemakers' Federation of Australia, 2017. National Vintage Report 2017 .

Winemakers' Federation of Australia, 2018. National Vintage Report 2018 .

Yegge, J.M., 2001. Influence of Sensory and Non-Sensory Attributes of Chardonnay Wine on Acceptance and Purchase Intent. Ph.D. thesis. ProQuest Dissertations Publishing.

ZHU, D.w., ZHANG, H.c., GUO, B.w., XU, K., DAI, Q.g., WEI, H.y., GAO, H., HU, Y.j., CUI, P.y., HUO, Z.y., 2017. Effects of nitrogen level on yield and quality of japonica soft super rice. Journal of Integrative Agriculture 16, 1018–1027. doi:10.1016/S2095-3119(16)61577-0.