Highlights

The influence of resource use on yield versus sale price trade-off in Australian vineyards

Author

- Comparative analysis of resource use, average sale price and quantity in Australian winegrowing.
- Regional comparison of outcomes and resource use in Australian winegrowing regions.
- Baseline models for comparing vineyards.
- Analysis of national, decade long data source.

The influence of resource use on yield versus sale price 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 sale price of the produce. In this analysis we observe relationships between resource use, yield and sale price through the use of statistical models. The dataset used for this analysis includes data collected for the past 10 years from 1261 vineyards located over a diverse range of Australian winegrowing regions. Yield and sale price were evaluated with respect to the resource use factors water use and Green House Gas (GHG) emissions. The analysis confirmed a strong relationship between area and resource use, with the overall area of a vineyard and its access to resources greatly determining the upper limit of yield. However, area was also negatively related to the average sale price of grapes; we find that higher average sale prices were connected to high resource inputs per area, rather than to the overall expenditure of resources. Regional and temporal effects on vineyard yield and average sales price were also identified. Overall, the analysis highlighted the importance of considering a vineyard's business goal, region, external pressures and economies of scale, when considering whether to pursue higher yields or higher average sales prices.

Preprint submitted to Journal of Agriculture and Food Research *March 12, 2024*

1. Introduction

The global focus on sustainability in agronomic industries has changed the way in which these enterprises do business. A dilemma exists across agriculture through shared fundamental considerations of resource use such as water and fuel, and the resultant yield and crop value that are produced (Hemming et al., 2020; Kawasaki and Uchida, 2016; ZHU et al., 2017). The average price of grapes for wine production is driven through its integration within the wine industry. An important connection between grapes and their price is the grape's perceived quality, which initially defines a wine's potential through the grape's chemical makeup (Black et al., 2015; Schreier and Jennings, 1979). Grapes of higher perceived quality or grapes from well known regions are likely to have higher prices (Wine Australia, 2021). Grape quality is connected to the market value of winegrapes, with Wine Australia explicitly defining grape quality through the use of discrete price brackets in their annual reports (Winemakers' Federation of Australia, 2018). It is important to note that the generalisation made to reflect quality through using average price assumes a due diligence of those who purchase the grapes (Yegge, 2001). The economic sustainability of a vineyard is tied to this market culture, driven by the wine industry. The consideration of sustainability within viticulture is also subject to environmental and sociodemographic pressures (Santiago-Brown et al., 2015). In the Australian context, these pressures include biosecurity, climate and international market changes (Canadell et al., 2021; Longbottom and Petrie, 2015; Oliver et al., 2013).

There is an extensive amount of research into the varied effects of factors on grape quality and yield (He et al., 2022; Laurent et al., 2022; Liakos et al., 2018). Due to the lack of long-term and in-depth data there is a lack of research on grape sale price and its driving factors. Furthermore, individual factors are often studied in isolation with yield and sales price not appearing together (Abbal et al., 2016). The lack of consolidated datasets restricts the ability to gain statistical insights at large scales and across multiple regions. As a result, broader studies are lacking (Keith Jones, 2002; Knight et al., 2019). The dataset used for this analysis includes data spanning 10 years from a multitude of vineyards located over a diverse range of Australian winegrowing regions. We use this dataset to describe the relationship of resources related to water and fuel use with the output yield and average sale price of the resultant product, taking into account the size and location of the vineyard. In a practical sense, this aim is a baseline for comparison: given a vineyard within Australia, one could estimate the comparative efficiency with regard to the trade-off between invested resources, yield and average sale price. 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 these analyses 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

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

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

statutory authority governed by the Wine Australia Act (Attorney-General's Department, 2010). Data collected by Wine Australia are publicly available. The response (predicted) variables in this analysis were yield, defined as the total tonnes of grapes harvested, and average sale price of grapes in Australian dollars per tonne. Both response variables were examined as totals and also relative to area harvested (see Table 1). Values were compared in this manner to observe how economies of scale affect the use of resources.

Data recorded by Sustainable Winegrowing Australia were entered manually by winegrowers voluntarily using a web based interface. For each model, vineyards were only included if they recorded all the variables used for the corresponding model (see Table 1). Each vineyard had at least recorded region, harvest year, yield and area harvested. Other variables that were used but not present for every single vineyard were 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. All variables were continuous except for harvest year and region, which were categorical variables (Table 1).

As data from Sustainable Wine Australia were voluntarily given, missing values were improved using regional average prices from the Wine Australia data. 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; with the data being publicly available.

The dataset was split into two distinct subsets that could be used across the different models. 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), as every sample/vineyard 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 discrete 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 (kg), electricity (Kw/h), and scope two emissions (TCO^2 equivalent). Variables considered but ultimately removed due to a lack of significant contributions to models, included the use of renewable energy, use of contractors, 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 of fuel to scope one emissions and prior calculations for all continuous variables which included logarithmic transformations, centring and scaling by standard deviation. We converted multiple emission sources into scope one emissions using the equation given from the Australian National Greenhouse Accounts Factors (AGDEE, 2021). The calculation conducted using

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

where emissions was the the product of the quantity of fuel in litres, Q, a prescribed Energy Content, EC, and emission factors (as given by the Australian National Greenhouse Accounts 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 defined by Wine Australia (Halliday, 2009; Oliver et al., 2013; SOAR et al., 2008). Although vineyards generally differed in income, there were pronounced differences between regions as shown in Figure 1. These differences were likely due to the site of a vineyard being important as it predetermines several physical parameters such as climate, geology and soil, making location a widely considered key determinant of grape yield and average sale price (Abbal et al., 2016; Agosta et al., 2012; Fraga et al., 2017). Each GI Region has its own unique mixture of climatic and geophysical properties that describes a unique winegrowing region within Australia and is a protected trademark under the Wine Australia act (Attorney-General's Department, 2010). Both Wine Australia and Sustainable Winegrowing Aus-

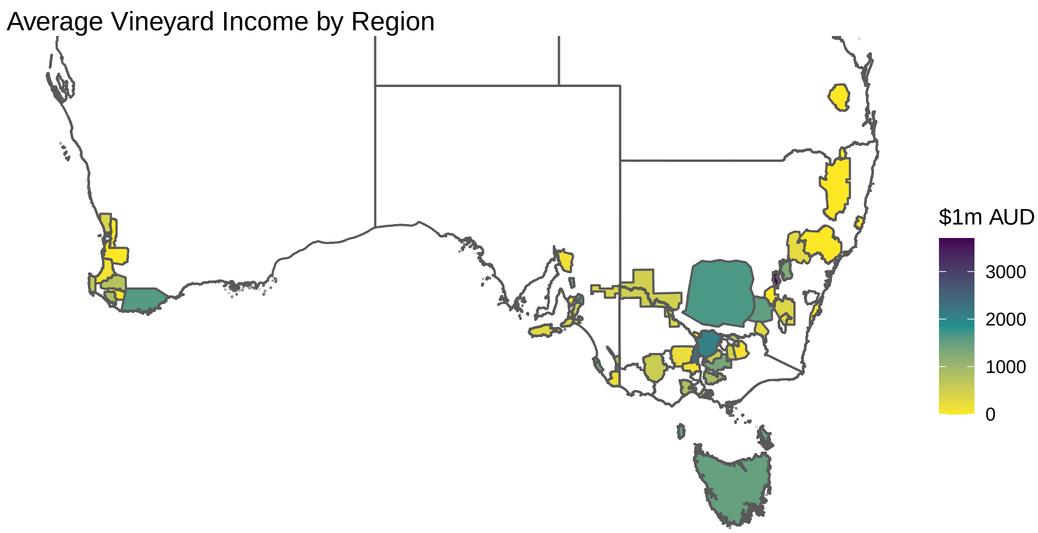

Figure 1: Map of each GI Regions' (outlined by Wine Australia) average income for a vineyard of that region (average grape sale price per tonne × total tonnes yielded).

tralia used the same GI Region categorical variable format to describe location.

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 (PPCC) 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, both the magnitude of the value and the associated confidence intervals were evaluated. P-values reflected the significance of a given correlation coefficient with statistical significance being declared when the corresponding P-value was lower than 0.05. PPCC analyses were undertaken 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 (GLMs) were created (see Table 1). GLMs 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 terms of R^2 and adjusted R^2 as well as F statistics, and t-tests were used to determine if predictors significantly contributed to the models when accounting for other variables, showing which specific years and areas contributed significantly. Both the PPCC and GLM analyses 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 little further insights or improvements in accuracy.

2.3. Model Validation

Models were validated using K-fold cross validation. This was performed by removing a subset of data from the sample used to train models and then predicting 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)	7.76E+02	2.18E+03	1.00E+00	7.23E+04
Area Harvested (ha)	6.67E+01	1.34E+02	7.00E-02	2.44E+03
Water Used (ML)	7.47E+06	5.65E+08	1.00E+00	4.27E+10
Scope One Emissions (tCO_2e)	4.17E+04	8.57E+04	6.76E+00	2.11E+06
$\frac{\text{Yield (tonnes)}}{\text{Area harvested (ha)}}$	1.01E+01	8.13E+00	4.00E-02	8.63E+01
Average Sale Price (AUD/tonne)	1.48E+03	9.22E+02	1.60E+02	2.60E+04
$\frac{\text{Average Sale Price (AUD/tonne)}}{\text{Area Harvested (ha)}}$	1.35E+02	5.71E+02	1.75E-01	2.98E+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).

PPCCs values for the transformed, centred and scaled variables are shown

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

in Table 3. The process of centring and scaling variables changes them so that their mean is centred on 0 and all values are scaled to the variables standard deviation, such that 1 is equal to the variables standard deviation. All correlations were found to be statistically significant ($p < 2.200E\text{-}16$), and except for Average Sale 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. The negative correlations between Average Sale Price and Yield, Water Use, Area and Scope One Emissions each indicated that Area Harvested and Fuel Use separately were not the determining factor for average sale price. The negative correlations are associative, not causal relationships (i.e using more water does not cause lower sale prices).

3.2. General Linear Models

Each model had a high R^2 value, indicating that most of the variance within the data was described by the models (see Table 4) and statistical

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

significance of F-tests ($p < 2.200E\text{-}16$). Aside from 3 variables, all regression parameters were also statistically significant from zero, F-tests across each model's variables were ($p < 0.05$). 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).

Coefficients related to continuous variables are summarised in Table 5. In Model 1 all coefficients except for the intercept significantly contributed to the model ($p < 0.05$), and in Model 2 all coefficients were statistically significant ($p < 0.05$). However, scope one emissions did not significantly contribute to models 3, 4 and 5. Model 4 only had 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 Model's coefficients for continuous variables

		Intercept	**Area Harvested**	**Water Used**	**Scope One Emissions**	**Area Harvested * Scope One Emissions**	**Area Harvested * Water Used**
Model 1	Coefficient	-3.32E-02	7.42E-01	8.66E-02	6.73E-02		
	Std Error	1.96E-02	1.00E-02	8.90E-03	8.00E-03		
Model 2	Coefficient	1.70E-01	5.77E-01	1.08E-01	8.50E-02	-4.97E-02	-5.35E-02
	Std Error	5.91E-02	1.48E-02	1.31E-02	1.17E-02	8.10E-03	8.40E-03
Model 3	Coefficient	1.81E-02	9.71E-01	-2.31E-02	-7.00E-03		
	Std Error	1.30E-02	7.20E-03	6.90E-03	5.70E-03		
Model 4	Coefficient	1.45E-01	2.40E-03	-4.66E-02	-1.70E-02	1.15E-02	1.40E-03
	Std Error	5.28E-02	1.50E-02	1.43E-02	1.18E-02	7.90E-03	8.30E-03
Model 5	Coefficient	1.52E-01		-4.04E-02	-1.71E-02		
	Std Error	5.27E-02		1.13E-02	9.70E-03		

highest value; instead, area harvested, being an order of magnitude greater, dominated the models. Model 5 achieved 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 variable of each model is depicted in Figure 2. For each model, the first year for a model's was used as the baseline: 2012 for Models 1 and 2, and 2015 for Models 3, 4 and 5. 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, the coefficients varied substantially over the years, and were not significantly different from zero in some years. Models 4 and 5 are very similar in Figure 2, 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 plateauing afterwards. Model 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 by temperature and rainfall are summarised in Figure 3. The most notable difference is between vineyards within 'Hot' and 'Very Dry' regions (warm inland regions), where high average sale prices are historically low, and yield is high. Water Use changes dramatically between these regions as well, with water being a driving force for yield but not necessarily average sale price. The warmer and drier regions tend to also have larger vineyards. These regional differences are further shown in Figure 4, as a ratio of vineyard; where there is a pronounced difference in 'Hot' and 'Very Dry' regions producing more per area with lower average sale prices per area than other regions.

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 as expected, indicating that the models are robust and consistent.

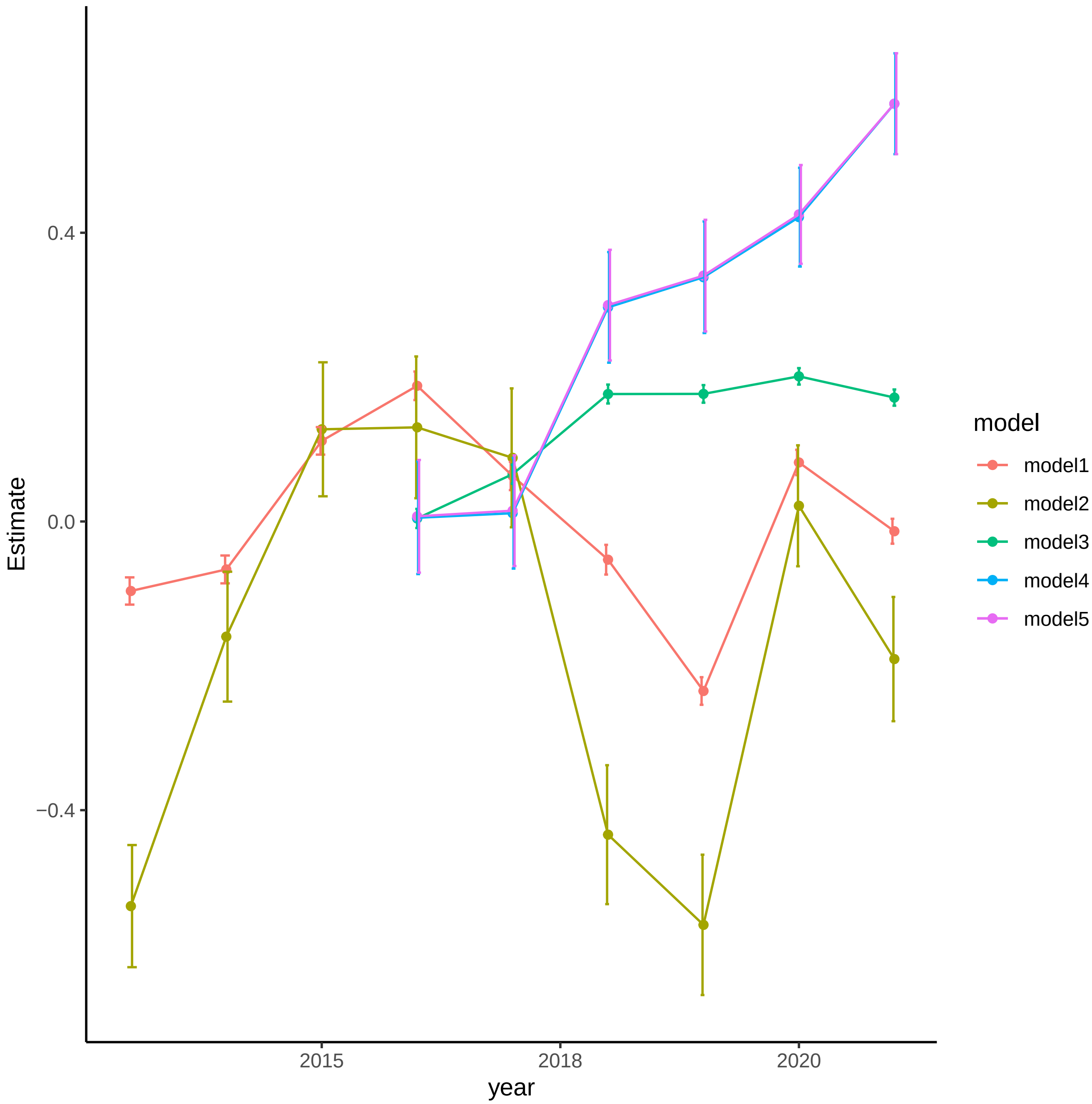

Figure 2: Model Coefficient values for Year, with standard error bars.

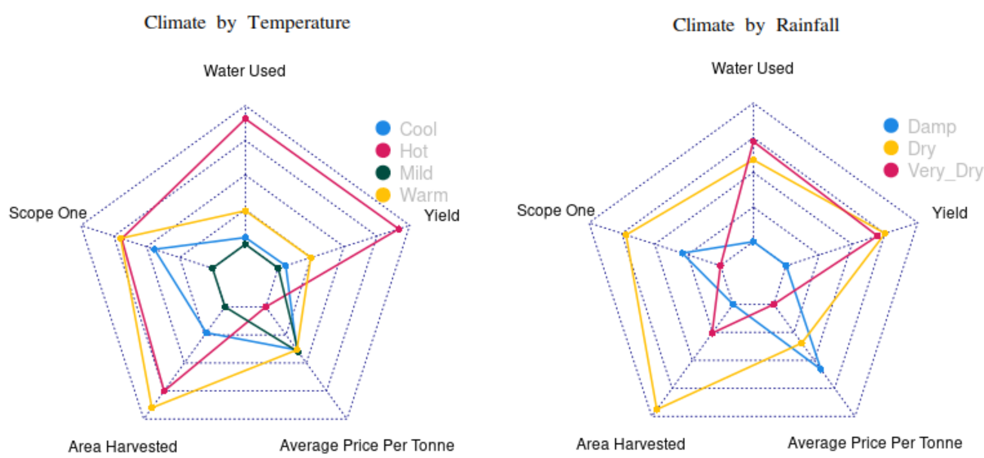

Figure 3: Radar plot of climatic profile's resource use, yield and average sale price. The left reflects vineyards in different climatic temperatures. The right reflects vineyards in different rainfall climates.

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

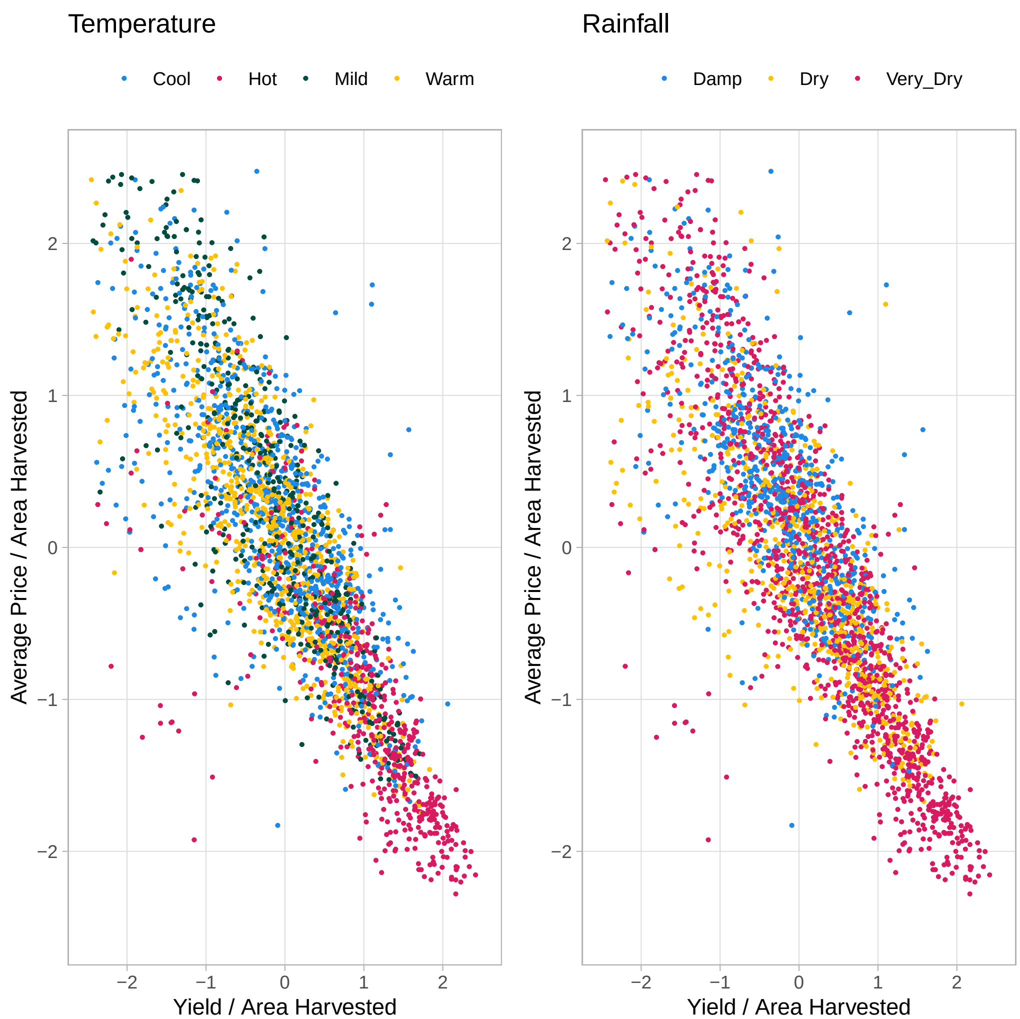

Figure 4: Scatter plot of vineyard yield against the average sale price as ratios to area harvested. The axes are in standard deviations with points coloured by climate.

4. Discussion

There was an expected strong relationship between size and resource use, with the overall area 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 average sale price. Higher average sales prices were also related to high resource inputs per area instead of the overall expenditure of resources. Vineyard yields and sales price changed greatly by region and year. Even given regional and yearly changes, there was a strong connection between smaller vineyards and higher sales prices. This could have been due to more attention available when managing smaller properties.

The lack of significance of scope one emissions and its contribution to models, given its F-statistics, could be indicative that other vineyard activities requiring fuel are not leading factors for a vineyards average sale price. 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 sales price and yield. The decision to prioritise average sale price 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 over the past decade (Wine Australia, 2019, 2021, 2022; Winemakers' Federation of Australia, 2013, 2014, 2015, 2016, 2017, 2018). 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 with lower average sales prices did not (Wine Australia, 2019, 2020, 2021; Winemakers' Federation of Australia, 2013, 2014, 2015, 2016, 2017, 2018). Vineyards in warm inland regions also tend to be larger, 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 an 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 effects, such as the use of cover crops, midrow crop rotation and efficient irrigation.

Some regions appeared to produce grapes of lower average sale price at scale whilst others achieved higher priced grapes in lower volumes. This empirical finding is consistent with Wine Australia's annual reports, which shows that some GI regions 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 only produce grapes of higher sales price but in smaller quantities. The difference in pricing per tonne between the lowest and highest regional average sales prices was almost a hundred times, showing that region had a profound influence. Some regions also had a mixture of high and low average sales price and yield showing regional variability in pricing which may be explained by varieties produced. A further possibility is the existence of regional upper limits on potential sales price, or that there are diminishing returns in some regions when pursuing higher sales prices or quantity; however these types of relationships may be obfuscated by knowledgeable winegrowers who avoid such pitfalls.

Due to regional differences, different strategies are employed, such as some regions targeting mass production over higher sales price. This is most notable when grouping regions by climate, especially when considering GI Regions in the 'Hot Very Dry' climate (see Figure 4). Figure 3 also shows that comparatively 'Warm' and 'Dry' Regions manage their resources incredibly efficiently having generally larger areas using similar resources to those in 'Cool' regions but having areas comparable to regions in 'Hot' climates. The coefficients for Model 4 also show a greater benefit of resource use per area when producing grapes at higher average sale prices, showing higher resource use per area reflective of higher average sales price (see Table 2). 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 sale price (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.

We identified two main limitations to our linear modelling. First, 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 fuel and water was likely used 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 fuel and water use 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.

The second limitations was the lack of further explanatory variables to help link models to causal affects. 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.

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 sales price 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 expected sale price and the quantity of grape yields. Average sales price of grapes was not solely tied to the overall expenditure of resources, but rather to the efficient allocation of resources per area. This emphasises that factors beyond sheer scale contribute significantly to the final sale price of grapes produced. Moreover, regional and yearly variations exhibited substantial effects on vineyard outputs, impacting both sales price and quantity. The connection observed between smaller vineyards and higher grape sale price suggests that the management of smaller properties might be more efficient and profitable.

References

Abad, J., Hermoso de Mendoza, I., Marín, D., Orcaray, L., Santesteban, L.G., 2021. Cover crops in viticulture. A systematic review (1):
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.

AGDEE, A.G.D.o.t.E.a.E., 2021. National Greenhouse Accounts Factors: 2021.

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.

Attorney-General's Department, 2010. Wine Australia Corporation Act 1980.

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

Black, C., Parker, M., Siebert, T., Capone, D., Francis, I., 2015. Terpenoids and their role in wine flavour: Recent advances. Australian Journal of Grape and Wine Research 21, 582–600. doi:10.1111/ajgw.12186.

Canadell, J.G., Meyer, C.P.M., Cook, G.D., Dowdy, A., Briggs, P.R., Knauer, J., Pepler, A., Haverd, V., 2021. Multi-decadal increase of forest burned area in Australia is linked to climate change. Nature Communications 12, 6921. doi:10.1038/s41467-021-27225-4.

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.

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.

Longbottom, M., Petrie, P., 2015. Role of vineyard practices in generating and mitigating greenhouse gas emissions. Australian Journal of Grape and Wine Research 21, 522–536. doi:10.1111/ajgw.12197.

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

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.

Santiago-Brown, I., Metcalfe, A., Jerram, C., Collins, C., 2015. Sustainability Assessment in Wine-Grape Growing in the New World: Economic, Environmental, and Social Indicators for Agricultural Businesses. Sustainability 7, 8178–8204. doi:10.3390/su7078178.

Schreier, P., Jennings, W.G., 1979. Flavor composition of wines: A review. C R C Critical Reviews in Food Science and Nutrition 12, 59–111. doi:10.1080/10408397909527273.

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.

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

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

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 .

Table A.7: P-values for the non-transformed water used variable's Pearson correlation coefficients.

Variable	Water Used
Yield	7.538E-01
Area	6.981E-01
Scope One Emissions	8.883E-01
$\frac{\text{Yield}}{\text{Area}}$	6.836E-01
Average Price Per Tonne	5.600E-02
$\frac{\text{Average Price per tonne}}{\text{Area}}$	1.522E-01

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.

Appendix A. Appendix

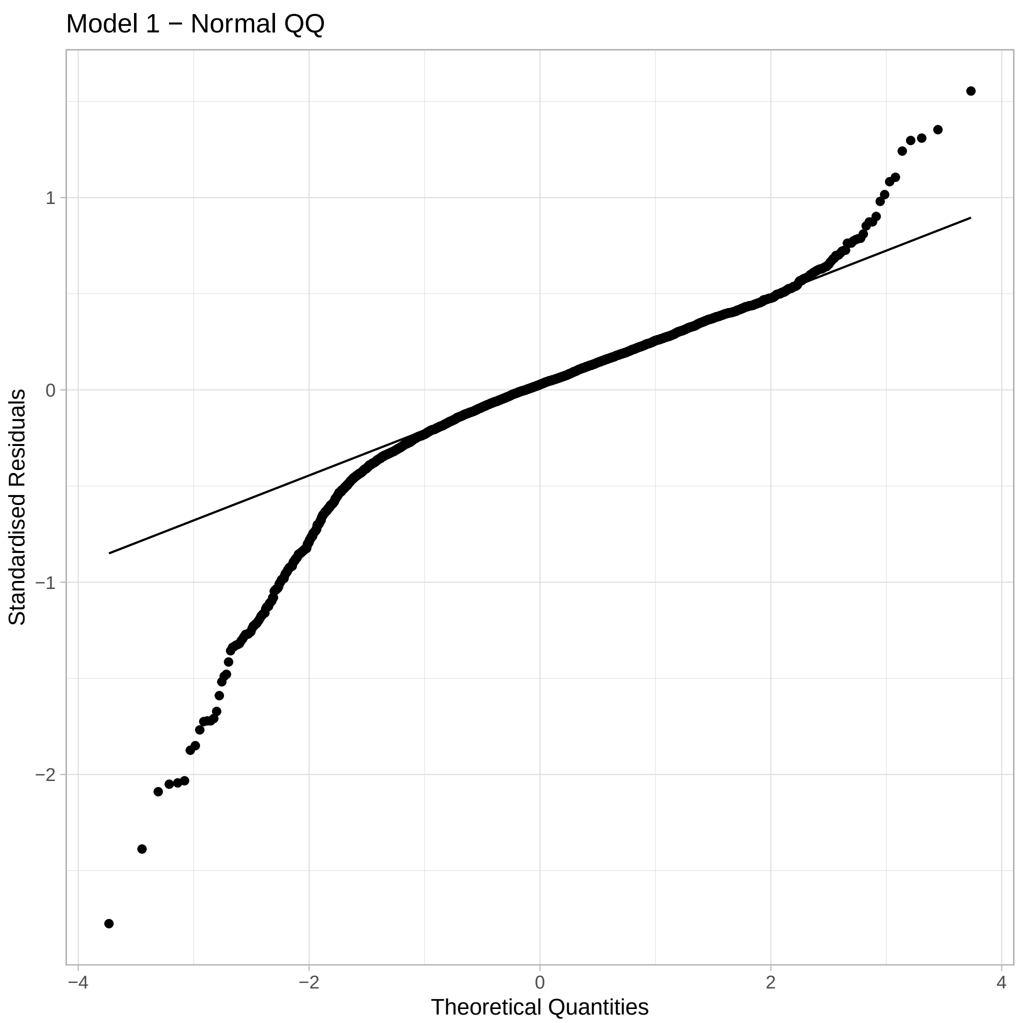

Figure A.5: QQ-plot of Model 1.

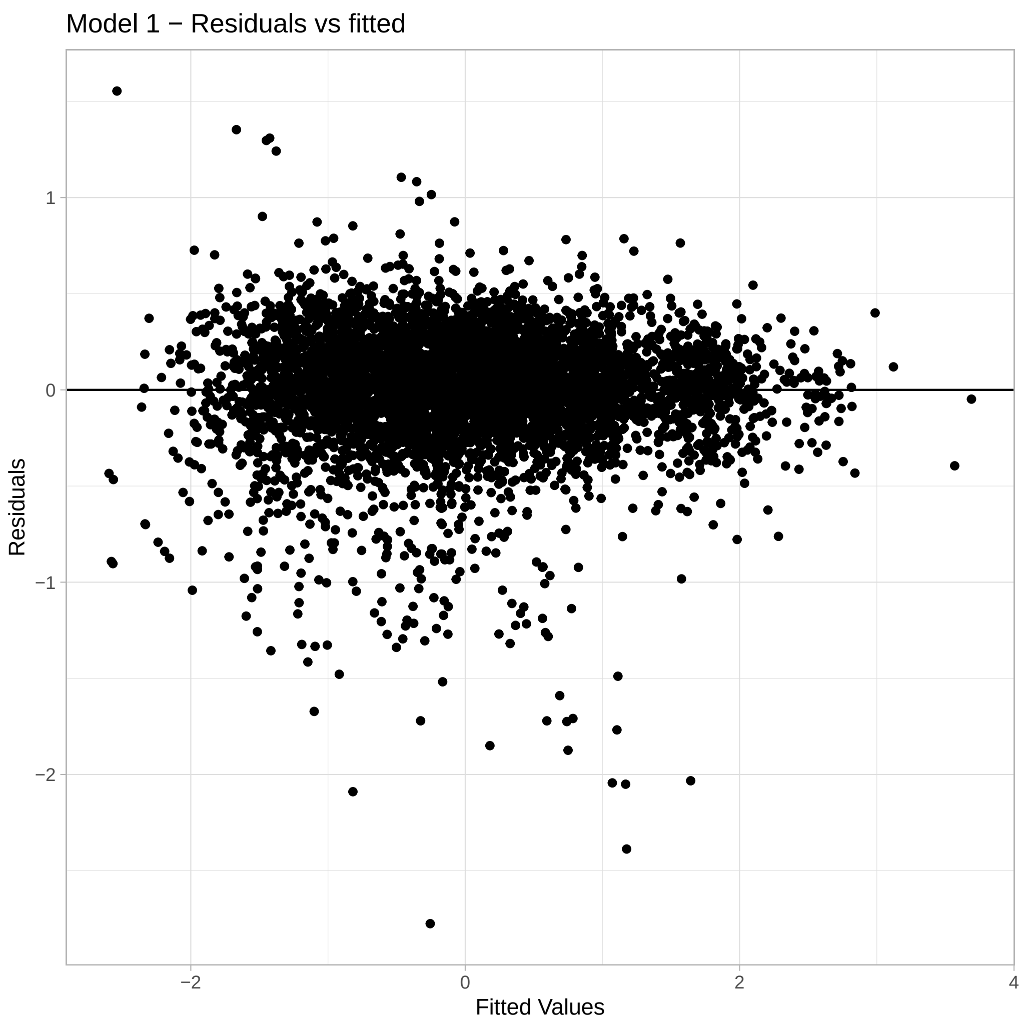

Figure A.6: Residuals vs fitted values for Model 1.

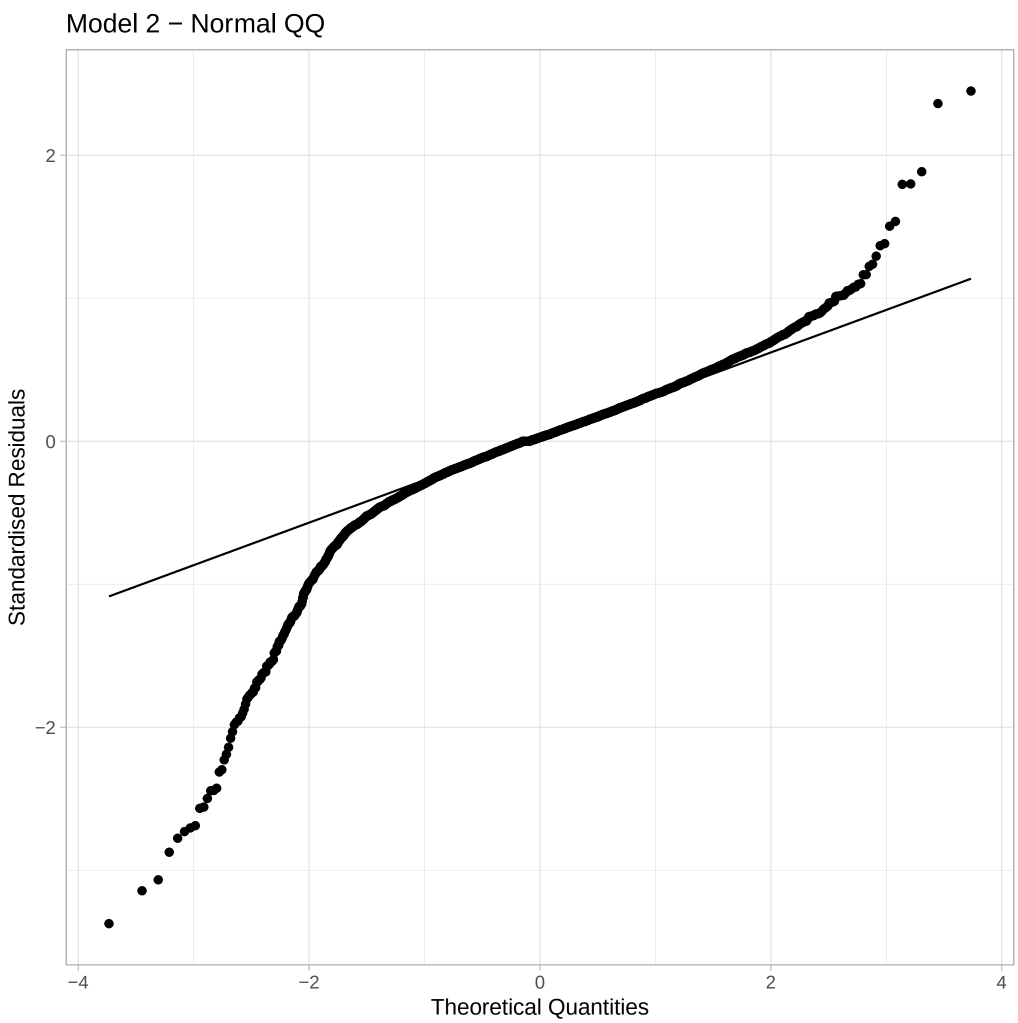

Figure A.7: QQ-plot of Model 2.

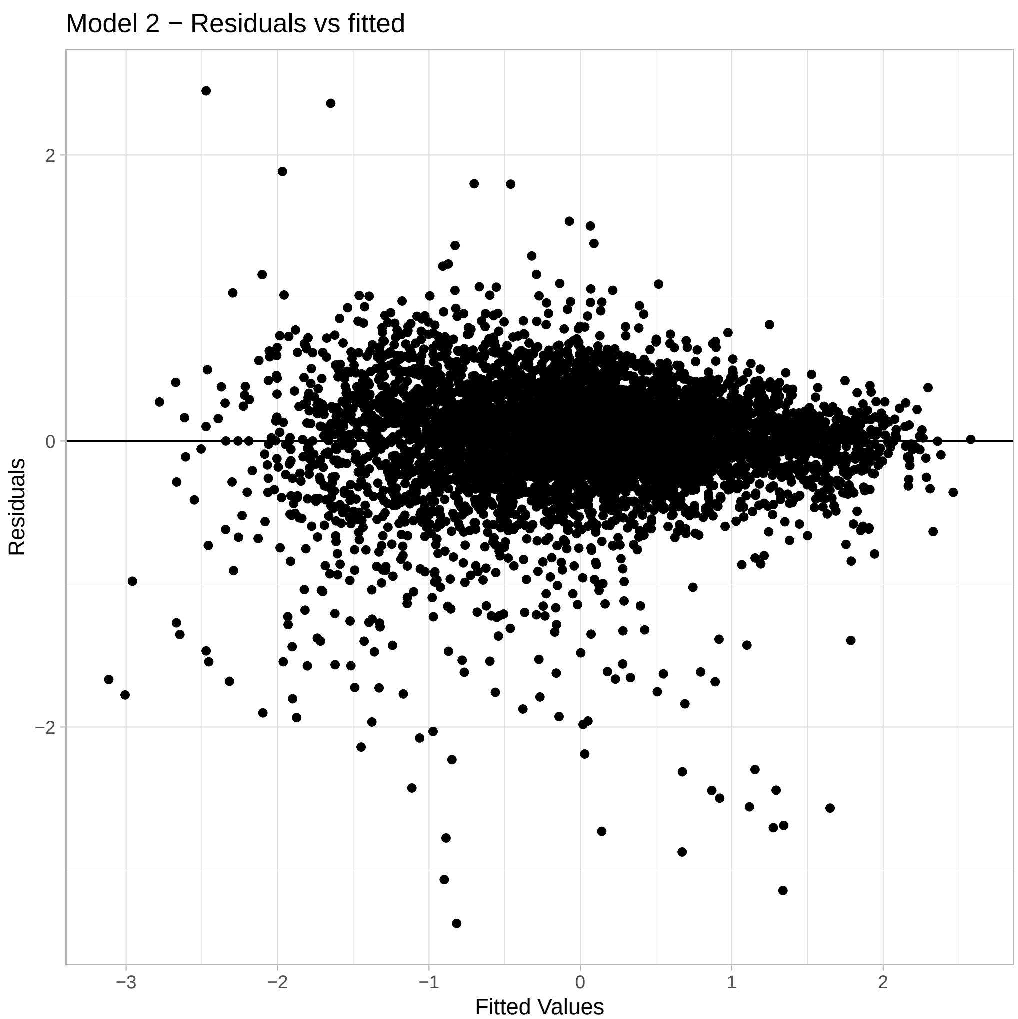

Figure A.8: Residuals vs fitted values for Model 2.

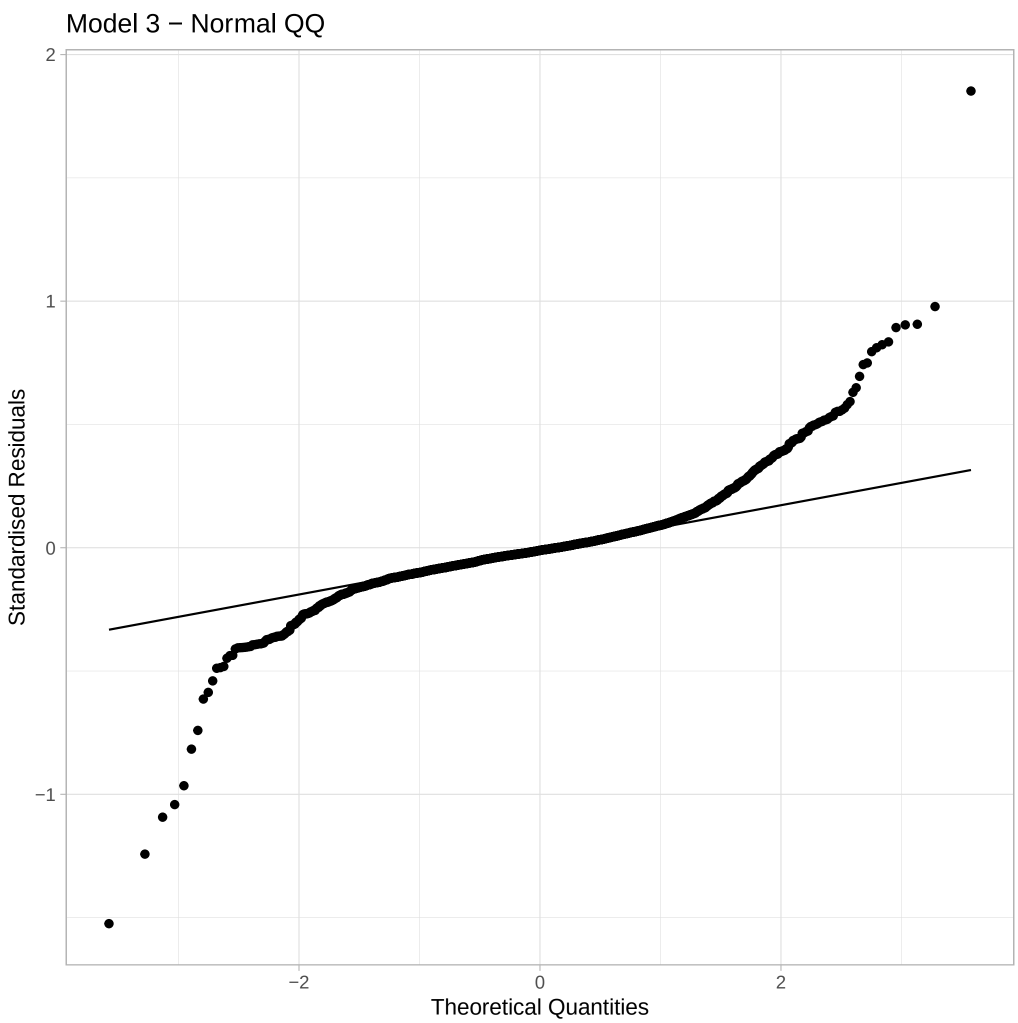

Figure A.9: QQ-plot of Model 3.

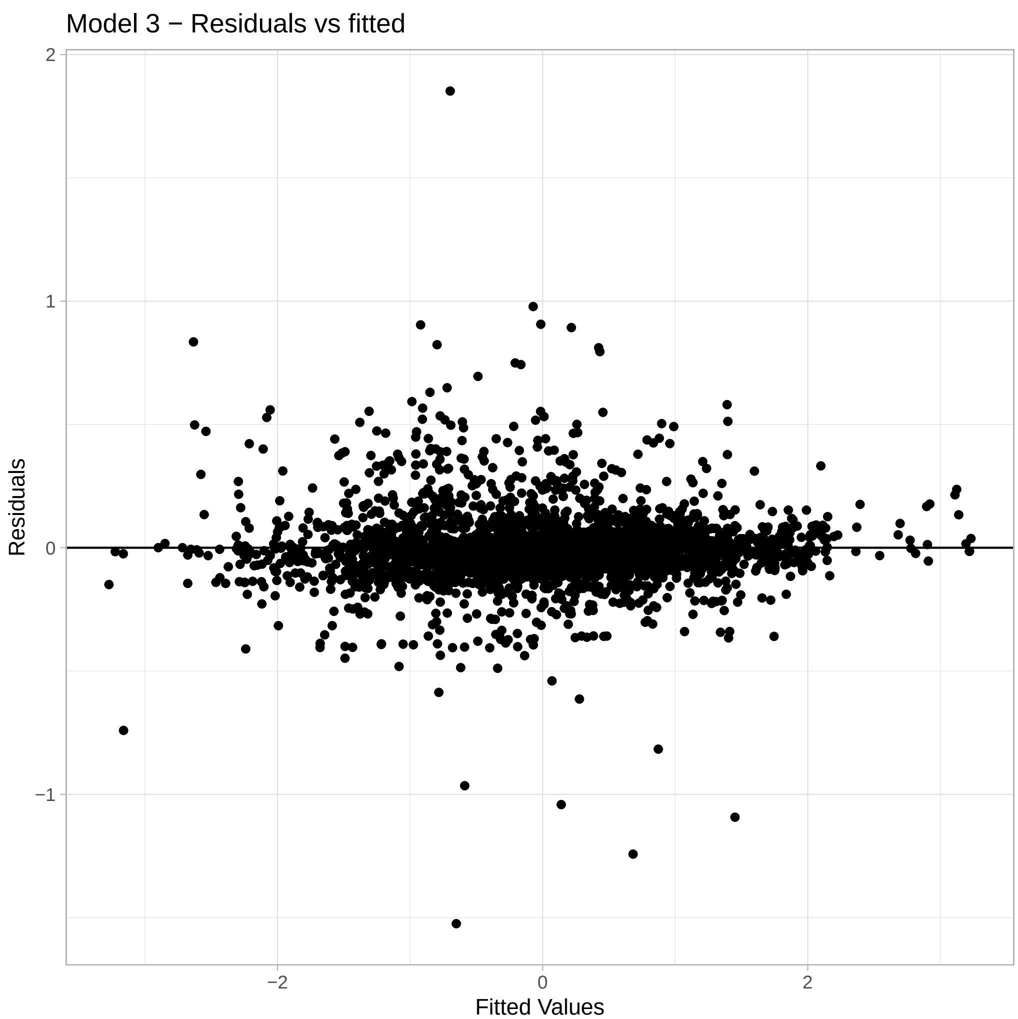

Figure A.10: Residuals vs fitted values for Model 3.

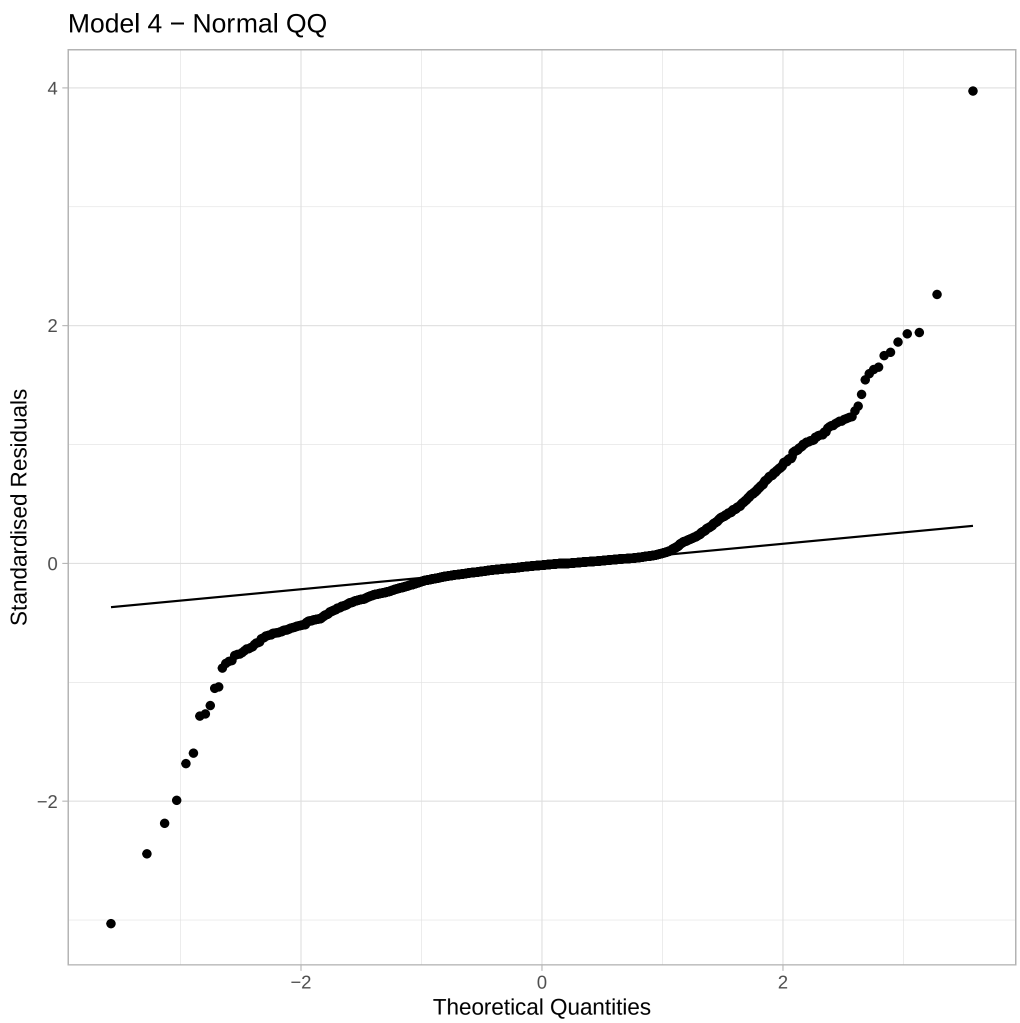

Figure A.11: QQ-plot of Model 4.

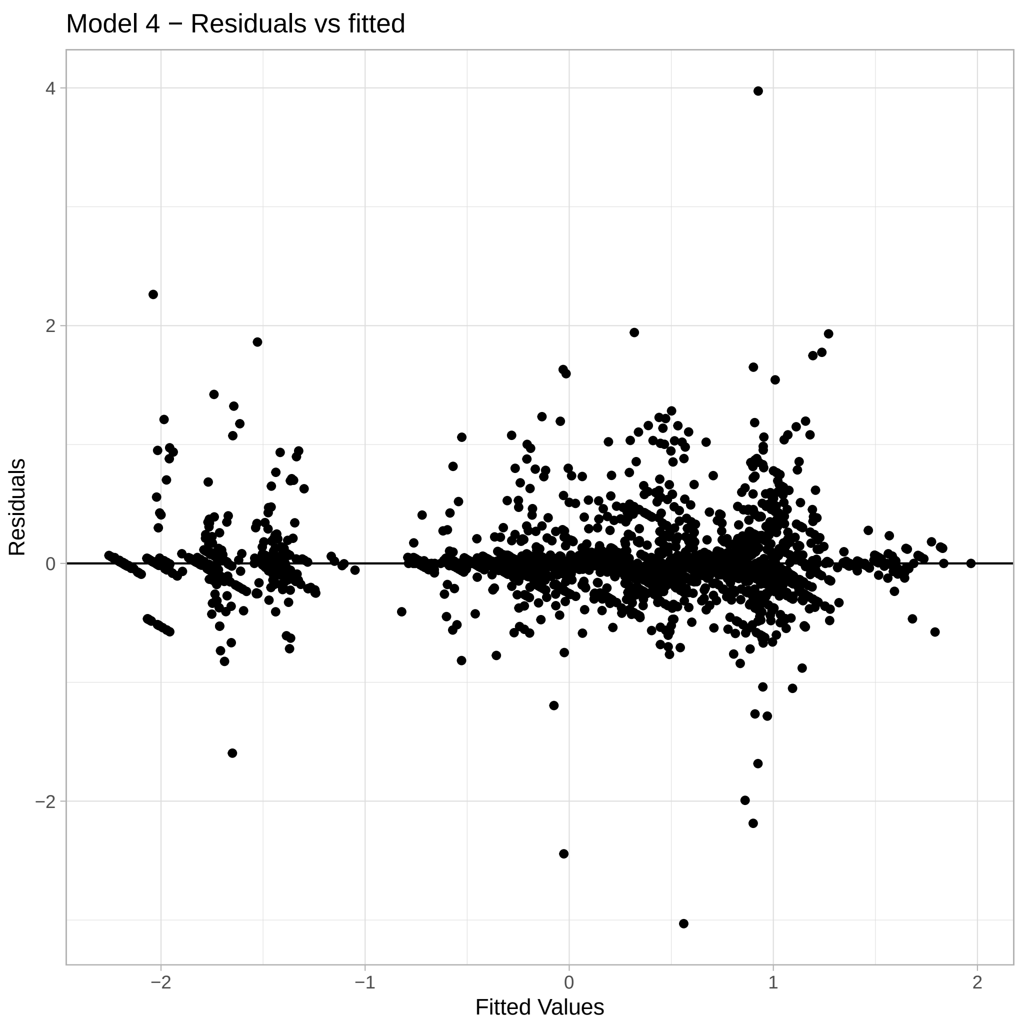

Figure A.12: Residuals vs fitted values for Model 4.

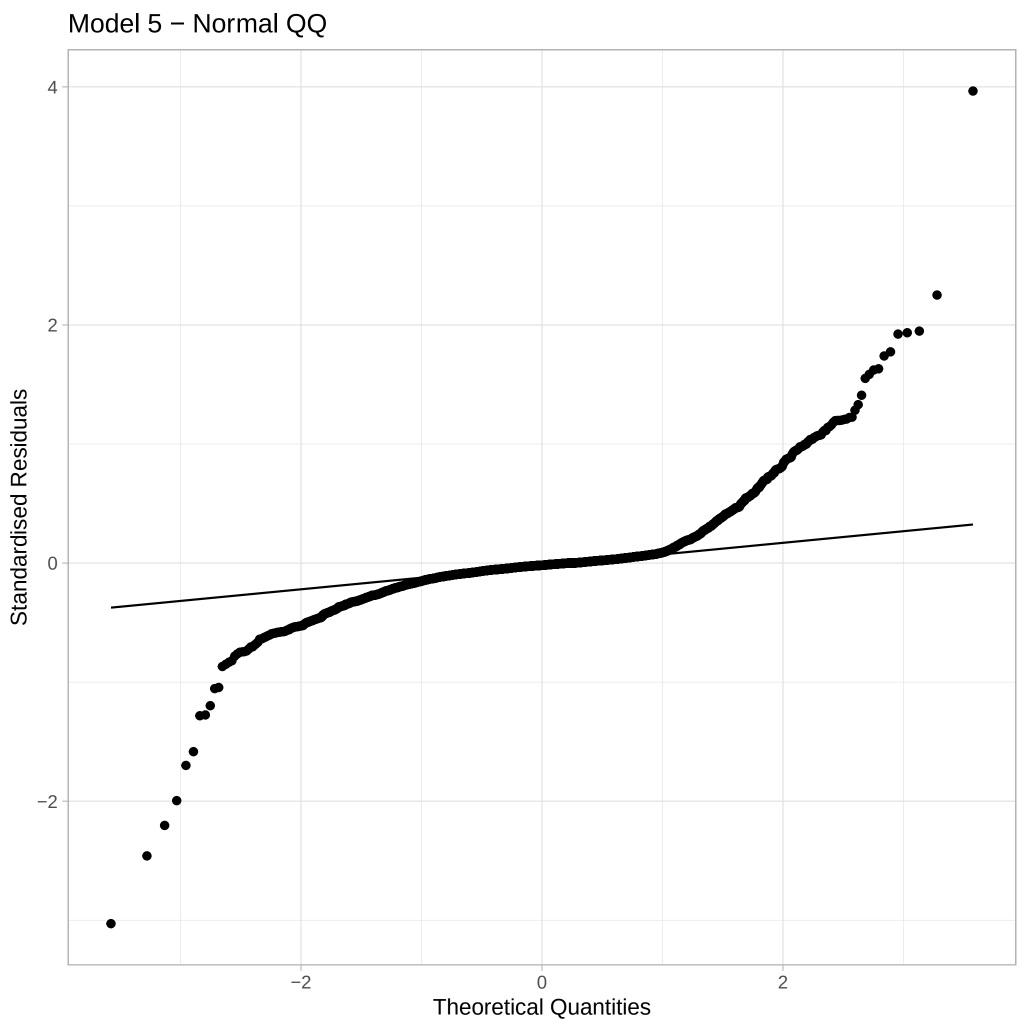

Figure A.13: QQ-plot of Model 5.

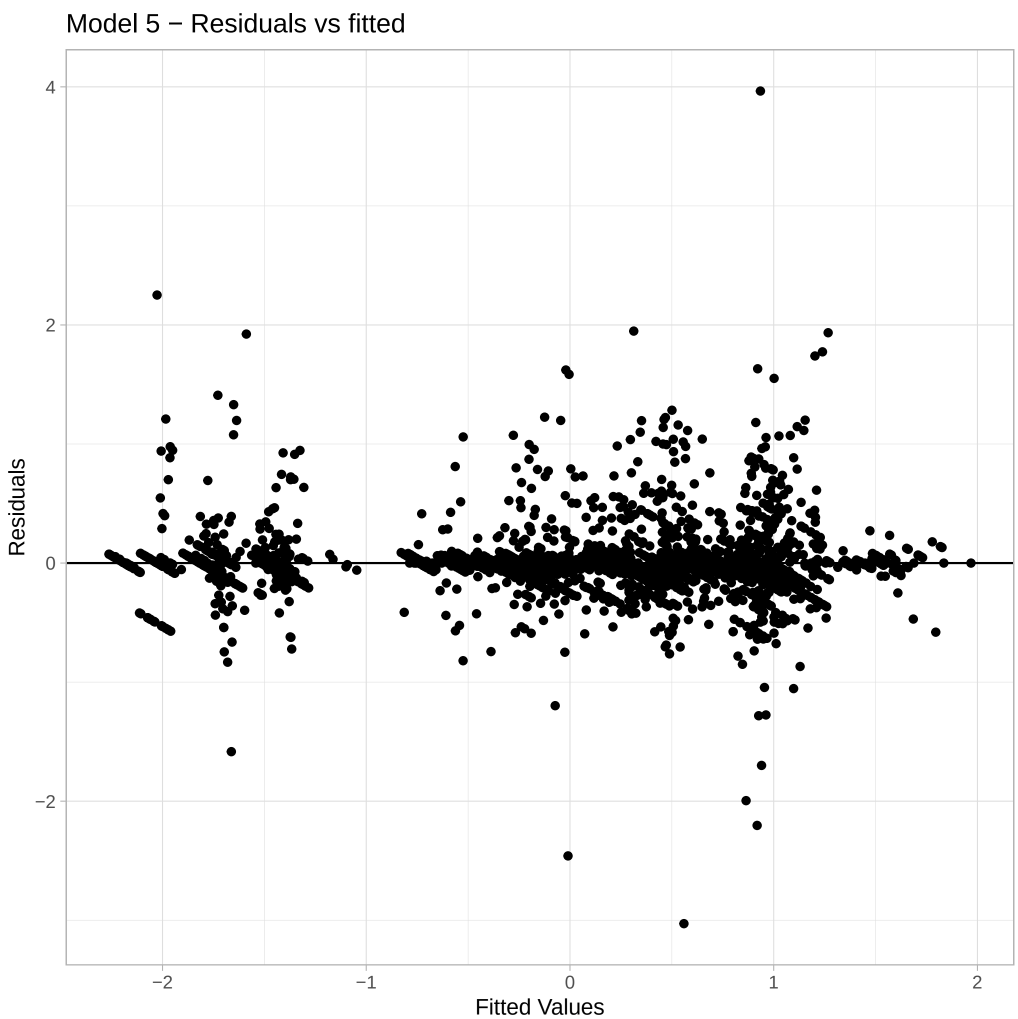

Figure A.14: Residuals vs fitted values for Model 5.