

¹ Highlights

² The influence of resource use on yield versus quality trade-off in
³ Australian vineyards

⁴ Author

- ⁵ ● Comparative analysis of resource use, quality and quantity in Aus-
⁶ tralian winegrowing.
- ⁷ ● Regional comparison of outcomes and resource use in Australian wine-
⁸ growing regions.
- ⁹ ● Baseline models for comparing wine crops.
- ¹⁰ ● Analysis of national, decade long data source.

¹¹ The influence of resource use on yield versus quality
¹² trade-off in Australian vineyards

¹³ Author^{1,1,1}

¹⁴ **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 1261 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 yield. However, size was also negatively related to the average sale price of grapes, we find that 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.

¹⁵ **1. Introduction**

¹⁶ The global focus on sustainability in agronomic industries has changed the
¹⁷ way in which these enterprises do business. When strategies for a sustainable
¹⁸ winegrowing industry are assessed, there is a trade-off between balancing the
¹⁹ amount of resources invested and the resultant yield versus quality produced.

²⁰ This dilemma exists across agriculture through shared fundamental consider-
²¹ ations such as water use and fuel usage (Hemming et al., 2020; Kawasaki and
²² Uchida, 2016; ZHU et al., 2017). Quality in viticulture (the cultivation of
²³ grapes for wine production) is driven through its integration within the wine
²⁴ industry, with the potential quality of a wine being initially defined through
²⁵ the chemical makeup of the grapes used in its production. The consideration
²⁶ of sustainability within viticulture is further complicated by environmental
²⁷ and socio-demographic pressures (Santiago-Brown et al., 2015). In the Aus-
²⁸ tralian context, these include biosecurity, climate and international market
²⁹ demands (Canadell et al., 2021; Longbottom and Petrie, 2015; Oliver et al.,
³⁰ 2013).

³¹ There is an extensive amount of research into the effects of a variety of
³² factors on grape quality and yield (He et al., 2022; Laurent et al., 2022; Liakos
³³ et al., 2018). However, due to the lack of long-term and in-depth data, indi-
³⁴ vidual factors are often studied in isolation (Abbal et al., 2016). The lack of
³⁵ consolidated datasets restricts the ability to gain statistical insights at large
³⁶ scales and across multiple regions, as a result broader studies are lacking
³⁷ (Keith Jones, 2002; Knight et al., 2019). The dataset used for this analysis

38 includes data spanning 10 years from a multitude of vineyards located over
39 a diverse range of Australian winegrowing regions. We use this dataset to
40 describe the relationship of resources related to water and fuel use with the
41 output yield and quality of the resultant product, taking into account the
42 size and location of the vineyard. The practical addition of this aim is a
43 baseline for comparison: given a vineyard within Australia, one could esti-
44 mate the comparative efficiency with regard to the tradeoff between invested
45 resources, yield and quality. This is the first time that such a trade off has
46 been confirmed explicitly across such varying regions, scales and climates in
47 the Australian winegrowing industry.

48 **2. Methods**

49 *2.1. Data*

50 Data used in this analysis were obtained from Sustainable Winegrow-
51 ing Australia and Wine Australia. Sustainable Winegrowing Australia is
52 Australia's national wine industry sustainability program, which aims to fa-
53 cilitate grape-growers and winemakers in demonstrating and improving their
54 sustainability (SWA, 2022). Wine Australia is an Australian Government
55 statutory authority governed by the Wine Australia Act (Attorney-General's
56 Department, 2010).

57 Predicted variables in this analysis were yield, defined as the total tonnes
58 of grapes harvested, and quality, defined as average sale price of grapes. It is
59 acknowledged that quality can be defined in a variety of ways, for example by
60 the grapes': aroma, chemical composition and color (Kasimati et al., 2022;
61 Mejean Perrot et al., 2022; Suarez et al., 2021). Using sale price was based

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	$\text{Yield} \times \text{Average Sale Price}$	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

62 on the reliance of market value of winegrapes on grape quality and because
63 Wine Australia explicitly defines grape quality through the use of discrete
64 price brackets in their annual reports. The generalisation made to reflect
65 quality through using average price assumed a due diligence of those who
66 purchased the grapes (Yegge, 2001). Both response variables were examined
67 as totals and as scales of area harvested. Values were compared in this
68 manner to observe how economies of scale affect the use of resources.

69 Data obtained from Wine Australia were collected via phone surveys and
70 included: total tonnes purchased, average price per tonne and yearly change
71 in price for region and grape varietal. Data recorded by Sustainable Wine-
72 growing Australia was entered manually by winegrowers using a web based
73 interface with some fields being optional. Required variables included: re-
74 gion, harvest year, yield and area harvested. Harvest year and region were
75 recorded as categorical variables. Optional variables included average sale
76 price, water used and fuel used (diesel, petrol, biodiesel and LPG). The
77 dataset was limited to respondents that recorded values for area harvested,
78 water used, yield and fuel use. To enable direct comparisons between fuels,
79 fuel use was converted to tonnes of Carbon Dioxide equivalent and collec-
80 tively referenced to as emissions.

81 Average sale price was an optional field in the Sustainable Winegrowing
82 Australia's dataset. Missing values were improved using regional average
83 prices from Wine Australia. Two subsets of data were then created for the
84 analysis. The first subset contained all vineyards and was used for two models
85 (Model 1 and Model 2, see Table 1). The second subset contained vineyards
86 which either recorded a value for average price of sale per tonne through

87 Sustainable Winegrowing Australia, or were within a region with an average
88 price of sale recorded by Wine Australia; this subset was used for three
89 further models (Models 3, 4 and 5, see Table 1). These subsets meant that
90 the data would be limited to samples which had recorded values for the
91 response variables (see Table 1), where every sample had a recorded value
92 for yield but not average price of sale per tonne.

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

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

101 Additional variables were considered for analysis but were excluded due to
102 being either underreported or had insignificant contributions to model accu-
103 racies. Variables explored but not used due to low reporting values included
104 fertiliser, and scope two emissions. Variables considered but ultimately re-
105 moved due to a lack of significant contributions to models, included the use
106 of renewable energy, contractor use, and pressures such as frost, fire and
107 disease.

108 Data preprocessing was conducted prior to analysis using the Python
109 programming language (G. van Rossum, 1995). Preprocessing included the
110 conversion from fuel to scope one emissions and prior calculations for all con-
111 tinuous variables which included logarithmic transformations, centring and

112 scaling by standard deviation. We converted multiple emission sources into
113 scope one emissions using the equation given from the Australian National
114 Greenhouse Accounts Factors (AGDEE, 2021), shown as

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

115 was used to convert the quantity of fuel in litres, Q , using a prescribed En-
116 ergy Content, EC , and emission factors of scope one, $EF1$, and scope three,
117 $EF3$, to tonnes of Carbon Dioxide Emission equivalent, $tCO2e$ (Department
118 of Climate Change, Energy, the Environment and Water, 2022).

119 The site of a vineyard predetermines several physical parameters such as
120 climate, geology and soil, making location a widely considered key determi-
121 nant of grape yield and quality (Abbal et al., 2016; Agosta et al., 2012; Fraga
122 et al., 2017). Differences in vineyard locations were captured through the use
123 of Geographical Indicator Regions (GI Regions, see Figure 1) defined by Wine
124 Australia (Halliday, 2009; Oliver et al., 2013; SOAR et al., 2008). Each GI
125 Region has its own unique mixture of climatic and geophysical properties that
126 describes a unique winegrowing region within Australia and is a protected
127 trademark under the Wine Australia act (Attorney-General's Department,
128 2010). Both Wine Australia and Sustainable Winegrowing Australia used
129 the same GI Region categorical variable format to describe location.

130 The climatic properties of each GI Region were summarised by using pre-
131 defined classifications as per the Sustainable Winegrowing Australia (2021)
132 user manual. The user manual describes climates by rainfall and tempera-
133 ture, creating supersets of Regions of similar climatic properties. The cli-
134 matic groups were used to illustrate similarities and differences occurring in

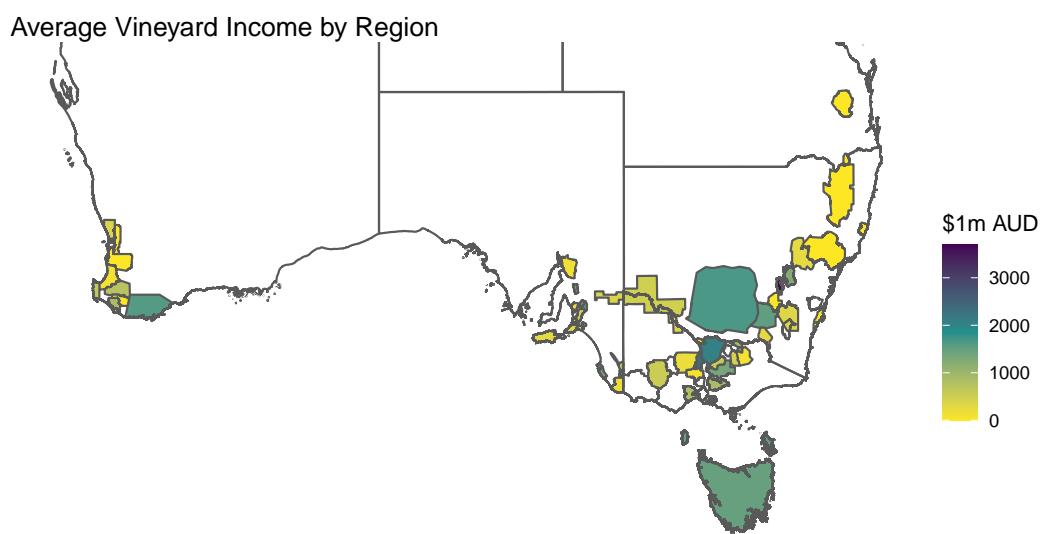


Figure 1: Map of vineyard average income for each of the used GI Regions.

135 areas larger than GI Regions.

136 *2.2. Analysis*

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

156 Five general linear models were created (see Table 1). Both the Pearson
157 Correlation Coefficients and General Linear Models were created using the
158 R statistical programming language (R Core Team, 2021) with the Caret
159 package (Kuhn, 2008). 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.

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.

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).

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

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

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

195 *3.2. General Linear Models*

196 Each model had a high R^2 value, indicating that a most of the variance
 197 within the data was described by the models (see Table 4). The models were
 198 found to be a good fit, with overall F-tests being statistically significant ($P <$

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

199 2.200E-16). And, aside from 3 variables, F-tests across each model's variables
 200 were also significant (with all being at least, $P < 0.05$). The three exceptions
 201 were: scope one emissions in Model 3 ($P=0.22$) and Model 4 ($P=0.0.39$), and
 202 the interaction between area harvested and water used in model 2 ($P=0.22$).
 203 Note that, scope one emissions was included in all models to directly compare
 204 the response variables as ratios of vineyard size to raw values and because
 205 it was strongly correlated to the response variable in every model (except
 206 model 5); especially for Models 1 and 4 (Table 3).

207 Models' continuous variable's coefficient values are summarised in Table
 208 5. Model 1 showed all coefficients except for the intercept were significantly
 209 contributing to the model ($P \leq 0.05$). Model 2's coefficients were all statis-
 210 tically significant. However, for Models 3, 4 and 5 Scope one emissions did
 211 not significantly contribute. And, Model 4 only saw statistically significant
 212 contributions from the intercept and water use. Although the coefficient for
 213 water use was statistically significant for each model, it did not have the
 214 highest value, instead area harvested, being an order of magnitude greater

Table 5: Summary of each Models coefficients for continuous variables

		Intercept	Area Harvested	Water Used	Scope One Emissions	Area Harvested	Area Harvested
						Scope One Emissions	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		

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

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

228 to 2018 but plateau afterwards. Models 1 and 2 do not show a clear trend
229 but do drop during 2017 and 2018 after increasing in the first 3 years.

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

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

241 Table 3.2 shows the validation results of each of the models. The R^2 mea-
242 sures of fit show similar results to the initial models, with a slight decrease.
243 Indicating that the models are robust and consistent.

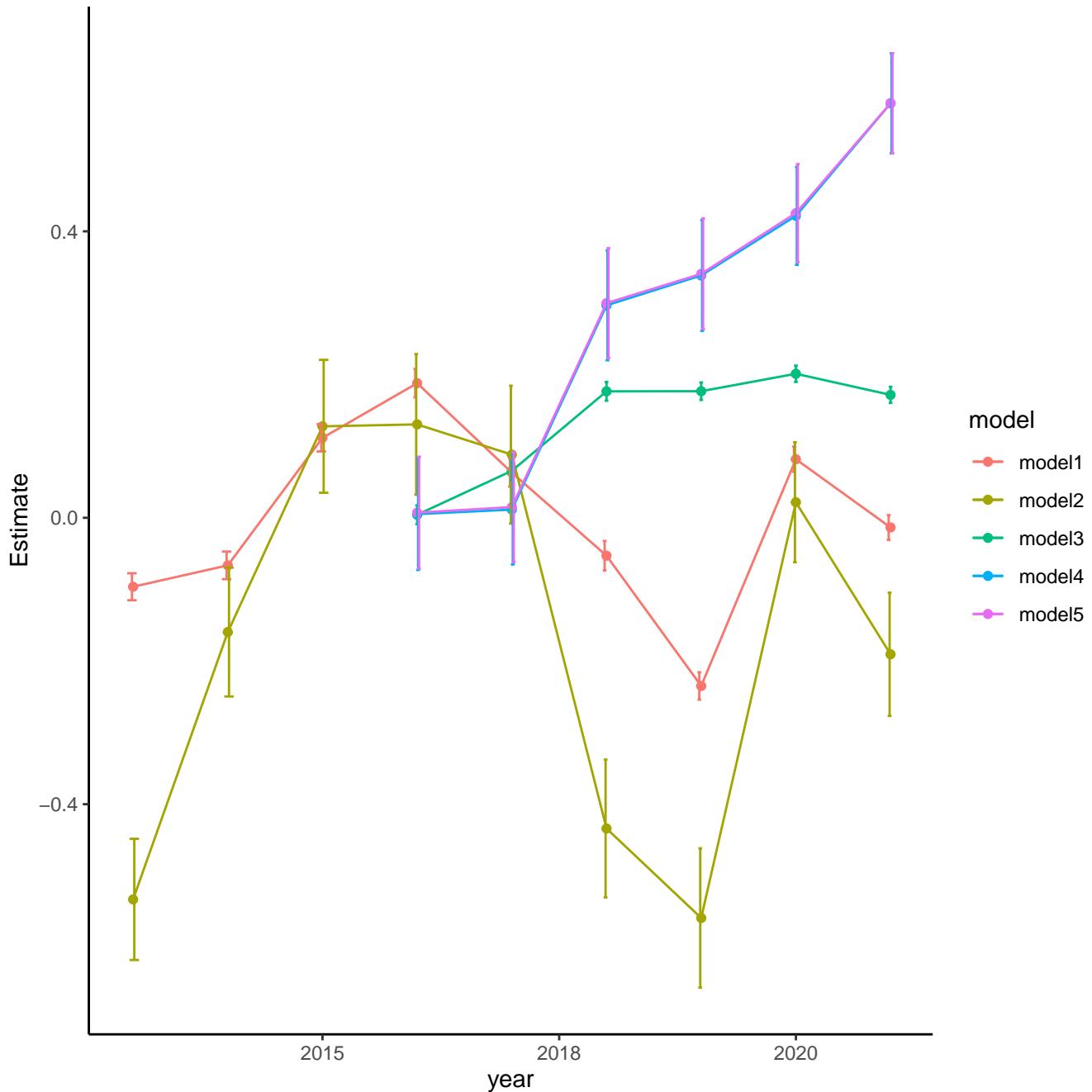


Figure 2: Model Coefficient values for Year, with standard error bars.

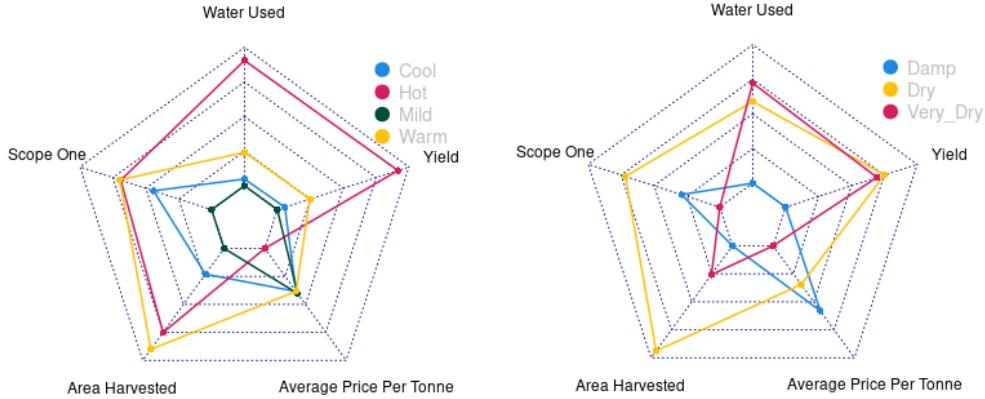


Figure 3: Radar plot of climatic profile's resource use, yield and average sale price.

244 4. Discussion

245 There was an expected strong relationship between size and resource use,
 246 with the overall space of a vineyard and its access to resources greatly deter-
 247 mining the upper limit of potential yield. However, size was also inversely
 248 related to the potential quality, with higher quality being related to high
 249 resource inputs per area; rather than to the overall expenditure of resources.
 250 Vineyard outputs were also augmented by regional and yearly affects. Even
 251 given regional and yearly changes, there was a strong connection between
 252 smaller vineyards and higher quality. This could have been due to the easier
 253 management of smaller properties.

254 Scope one emissions' lack of significance and contribution given its F-
 255 statistics, could be indicative that other vineyard activities requiring fuel are
 256 not leading factors for a vineyards grape quality. The relationship between
 257 yield, value and area was not simply about efficiently producing the most
 258 grapes. It is possible that the relationship of scope one emissions between

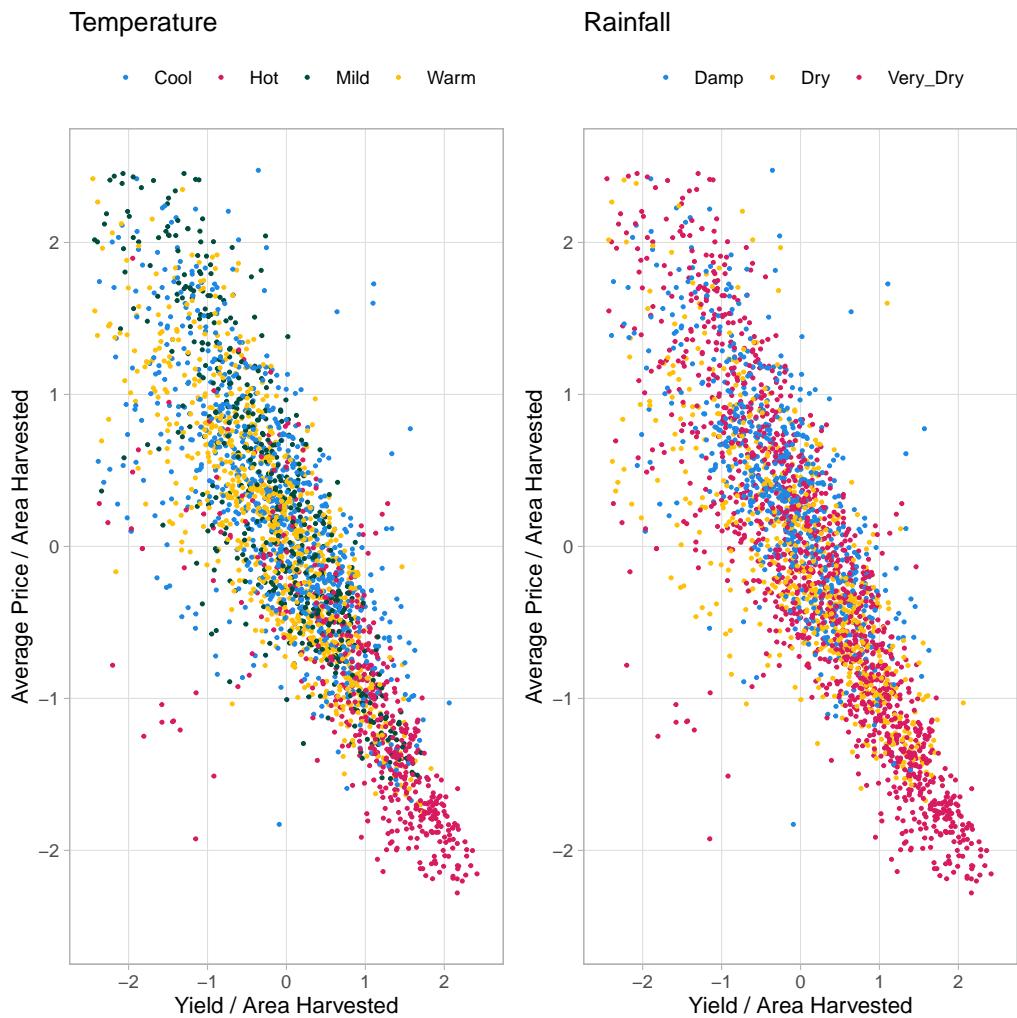


Figure 4: Scatter plot of vineyard yield against the average sale price as ratios to area harvested. The axes are in standard deviations with points coloured by climate.

259 yield and sale price was closely tied to a vineyard's area due to requiring more
260 fuel to address more issues over greater distances. It is difficult to discern the
261 connection of scope one emissions directly, as fuel can be used for a broad
262 category of activities.

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

275 There are many on-the-ground decisions that influence both quality and
276 yield. The decision to prioritise quality over quantity, is governed by com-
277 plex physical and social forces, for example international market demands,
278 disease pressures and natural disasters (Abad et al., 2021; Cortez et al., 2009;
279 Hall et al., 2011; I. Goodwin, et al., 2009; Kasimati et al., 2022; Oliver et al.,
280 2013; Srivastava and Sadistap, 2018), with many of these occurrences being
281 highlighted throughout the reports from Wine Australia (Wine Australia,
282 2019, 2021, 2022; Winemakers' Federation of Australia, 2013, 2014, 2015,
283 2016, 2017, 2018) over the past decade. However, the changes in the coef-

284 ficients (see Figure 2) are not reflective of many known occurrences, such
285 as the 2020 bush fires, which had higher values for coefficients than prior
286 years; During the 2020 bush fires 40,000 tonnes of grapes were lost across 18
287 different wine regions due to bush fires and smoke taint. In comparison to
288 countrywide pressures such as drought, this damage made up only 3% of the
289 total amount of grapes for that year; although acknowledged as a consider-
290 able loss on an individual basis, it was deemed to be only a minor national
291 concern by Wine Australia when compared to other environmental pressures
292 such as drought (Wine Australia, 2020)

293 Climatic pressures are an important consideration for growers, especially
294 those in warmer and drier regions. The Wine Australia reports also show
295 that warm inland regions have seen a decline in profit over the past decade,
296 whereas regions targeting quality did not. The warm inland regions also
297 tend to contain larger vineyards, making up for lower sale prices with larger
298 yields. Considering the negative correlation of average price to area, for this
299 strategy to work economies of scale become an important factor. Given the
300 large quantities of grapes that can be produced by some vineyards, even at
301 low margins there is the potential to be profitable. However, the increasing
302 climatic pressures mixed with the requirement for larger volumes of water,
303 make the sustainability of some vineyards come into question. Furthermore,
304 intensive farming in general is known to jeopardise the sustainability of an
305 operation through the degradation of soil and waterways (Capello et al.,
306 2019; Lin, 2012; Pisciotta et al., 2015). There are established methods that
307 can help to mitigate these affects, such as the use of cover crops and crop
308 rotation. However, it has become more apparent that the active reduction of

309 grape yield, through methods such as thinning, can help increase the quality
310 of grapes and improve soil health (Condurso et al., 2016; Wang et al., 2019).

311 Some regions appeared to produce many low quality grapes at scale whilst
312 others focussed on producing higher quality grapes in lower volumes. This
313 empirical finding is consistent with Wine Australia's annual reports, which
314 shows that some GI regions, such as the Riverland, are known for producing
315 large amounts of lower grade (low value per tonne) grapes (Wine Australia,
316 2022; Winemakers' Federation of Australia, 2017). Comparatively, other re-
317 gions, such as Tasmania, only produce high quality grapes but in smaller
318 quantities. The difference in pricing per tonne between the lowest and high-
319 est graded grapes can be greater than a hundred times the difference in
320 value per tonne. Not all regions target only one grade of grape, with some
321 producing a variety of differently graded grapes; such as the Yarra Valley,
322 which produces grades from C to A. This effect could stifle the potential
323 for market opportunities within lesser known regions. A further possibility
324 is the existence of regional upper limits on potential quality, or that there
325 are diminishing returns in some regions when pursuing quality or quantity;
326 however these types of relationships may be obfuscated by knowledgeable
327 winegrowers who avoid this pitfall.

328 Due to regional differences, different strategies are also employed across
329 different regions, such as some regions targeting mass production over quality.
330 This is most notable when grouping regions by climate, especially when con-
331 sidering GI Regions in the 'Hot Very Dry' climate (see Figure 4). Although
332 not chosen over GI region, climate was considered to be a large determinant
333 of the ability to produce larger quantities of grapes, as well as a determinant

334 in grape quality (Agosta et al., 2012). The more granular GI Region likely
335 explained a broader mix of geographical phenomenon, such as soil, geology
336 and access to water resources (Abbal et al., 2016; Carmona et al., 2011). The
337 interaction between year and GI Region likely accounted for events such as
338 bushfires, which would be impactful, but only at a local level, both in time
339 and space.

340 We identified two main limitations to our linear modeling. First model
341 1 and 2 over-predicting yield may have been due to preventative measures
342 brought on by regional pressures such as fire, frost and disease. More re-
343 sources were required to prevent these issues from spreading within a region,
344 thus disproportionately affecting some vineyards compared to others locally.
345 This type of maintenance is not well captured in the models, especially when
346 considering that some regions, especially those in warmer areas, are not as
347 prone to disease as cooler climates and could potentially have lower operating
348 costs per hectare. This could create a discrepancy in vineyards that utilised
349 preventative measures in wetter regions, as opposed to those that did not,
350 thus expending less fuel and energy but risking disease. When reviewing
351 the differences between regions, it is important to consider that vineyards in
352 'Hot Very Dry' areas can be hundreds of times the size of those in other re-
353 gions. This limitation could be overcome by incorporating the profitability of
354 vineyards, comparing the financial success of working at different operational
355 scales.

356 The second limitation was the lack of further explanatory variables to
357 help link models to causal affects. Variables such as the utilisation of renew-
358 able energy, contractors, and the occurrence of disease, fire and frost were

359 originally explored to capture the discrepancies between similar vineyards
360 that produced different yields and crop values. However, none of these vari-
361 ables was significantly correlated with the response variables, and did not
362 add to model accuracy, even when considered as interactions. Allowance
363 for nonlinear relationships, specifically through splines, resulted in more nor-
364 mally distributed residuals but at a drastically reduced overall accuracy when
365 comparing R^2 and Residual Square Error. Attempts to fully explain small
366 variations was always overshadowed by the dramatic differences in regional
367 trends. Having more data for each region would also be beneficial, allowing
368 greater comparison between regions. More variables may also help to discern
369 vineyards that can produce larger volumes of grapes at higher prices.

370 The use of other models such as random forests and decision trees along-
371 side more variables and data may help to uncover the reasons for under or
372 overestimation. These differences could be caused by the use of alternative
373 sustainable practices in the field. Moreover, while there is evidence to suggest
374 that environmentally sustainable practices can reduce costs, and increase effi-
375 ciency whilst improving the quality of grapes; more research is needed to link
376 these benefits across different regions and climates (Baiano, 2021; Mariani
377 and Vastola, 2015; Montalvo-Falcón et al., 2023).

378 5. Conclusion

379 This study delved into the relationships between resource use, grape qual-
380 ity and yield. The findings underscore the multifaceted nature of vineyard
381 management, where the interplay of size, resource allocation, climate, and
382 regional influences collectively shape both the quality and quantity of grape

383 yields. Quality was not solely tied to the overall expenditure of resources, but
384 rather to the efficient allocation of resources per unit area. This emphasises
385 that factors beyond sheer scale contribute significantly to the final quality
386 of the grapes produced. Moreover, regional and yearly variations exhibited
387 substantial effects on vineyard outputs, impacting both quality and quan-
388 tity. The connection observed between smaller vineyards and higher grape
389 quality suggests that the management of smaller properties might be more
390 streamlined and effective, enabling a greater quality of grape to be produced.

391 **References**

- 392 Abad, J., Hermoso de Mendoza, I., Marín, D., Orcaray, L., Santeste-
393 ban, L.G., 2021. Cover crops in viticulture. A systematic review (1):
394
Implications on soil characteristics and biodiversity in vineyard.
395 OENO One 55, 295–312. doi:10.20870/oeno-one.2021.55.1.3599.
- 396 Abbal, P., Sablayrolles, J.M., Matzner-Lober, É., Boursiquot, J.M., Baudrit,
397 C., Carbonneau, A., 2016. Decision Support System for Vine Growers
398 Based on a Bayesian Network. Journal of agricultural, biological, and
399 environmental statistics 21, 131–151. doi:10.1007/s13253-015-0233-2.
- 400 AGDEE, A.G.D.o.t.E.a.E., 2021. National Greenhouse Accounts Factors:
401 2021.
- 402 Agosta, E., Canziani, P., Cavagnaro, M., 2012. Regional climate variability
403 impacts on the annual grape yield in Mendoza, Argentina. Journal of
404 Applied Meteorology and Climatology 51, 993–1009.

- 405 Attorney-General's Department, 2010. Wine Australia Corporation Act
406 1980.
- 407 Baiano, A., 2021. An Overview on Sustainability in the Wine Production
408 Chain. *Beverages* 7. doi:10.3390/beverages7010015.
- 409 Canadell, J.G., Meyer, C.P.M., Cook, G.D., Dowdy, A., Briggs, P.R.,
410 Knauer, J., Pepler, A., Haverd, V., 2021. Multi-decadal increase of forest
411 burned area in Australia is linked to climate change. *Nature Communications*
412 12, 6921. doi:10.1038/s41467-021-27225-4.
- 413 Capello, G., Biddoccu, M., Ferraris, S., Cavallo, E., 2019. Effects of Tractor
414 Passes on Hydrological and Soil Erosion Processes in Tilled and Grassed
415 Vineyards. *Water* 11. doi:10.3390/w11102118.
- 416 Carmona, G., Varela-Ortega, C., Bromley, J., 2011. The Use of Participatory
417 Object-Oriented Bayesian Networks and Agro-Economic Models for
418 Groundwater Management in Spain. *Water resources management* 25,
419 1509–1524. doi:10.1007/s11269-010-9757-y.
- 420 Condurso, C., Cincotta, F., Tripodi, G., Sparacio, A., Giglio, D.M.L., Sparla,
421 S., Verzera, A., 2016. Effects of cluster thinning on wine quality of Syrah
422 cultivar (*Vitis vinifera* L.). *European food research & technology* 242,
423 1719–1726. doi:10.1007/s00217-016-2671-7.
- 424 Cortez, P., Teixeira, J., Cerdeira, A., Almeida, F., Matos, T., Reis, J., 2009.
425 Using data mining for wine quality assessment, in: *Discovery Science: 12th
426 International Conference, DS 2009, Porto, Portugal, October 3-5, 2009* 12,
427 Springer. pp. 66–79.

- 428 Department of Climate Change, Energy, the Environment and Water, 2022.
- 429 Australian National Greenhouse Accounts Factors.
- 430 Fraga, H., Costa, R., Santos, J.A., 2017. Multivariate clustering of viticultural
431 terroirs in the Douro winemaking region. Ciênc Téc. Vitiv. 32,
432 142–153.
- 433 G. van Rossum, 1995. Python tutorial, Technical Report CS-R9526. Centrum
434 voor Wiskunde en Informatica (CWI),.
- 435 Hall, A., Lamb, D.W., Holzapfel, B.P., Louis, J.P., 2011. Within-season
436 temporal variation in correlations between vineyard canopy and winegrape
437 composition and yield. Precision Agriculture 12, 103–117.
- 438 Halliday, J.C.J.C., 2009. Australian Wine Encyclopedia. Hardie Grant
439 Books, VIC.
- 440 He, L., Fang, W., Zhao, G., Wu, Z., Fu, L., Li, R., Majeed, Y.,
441 Dhupia, J., 2022. Fruit yield prediction and estimation in orchards:
442 A state-of-the-art comprehensive review for both direct and indirect
443 methods. Computers and Electronics in Agriculture 195, 106812.
444 doi:10.1016/j.compag.2022.106812.
- 445 Hemming, S., de Zwart, F., Elings, A., Petropoulou, A., Righini, I., 2020.
446 Cherry tomato production in intelligent greenhouses-sensors and ai for con-
447 trol of climate, irrigation, crop yield, and quality. Sensors (Basel, Switzer-
448 land) 20, 1–30. doi:10.3390/s20226430.
- 449 I. Goodwin,, L. McClymont,, D. Lanyon, A. Zerihun, J. Hornbuckle, M.

- 450 Gibberd, D. Mowat, D. Smith, M. Barnes, R. Correll, 2009. Managing soil
451 and water to target quality and reduce environmental impact.
- 452 Kasimati, A., Espejo-García, B., Darra, N., Fountas, S., 2022. Predicting
453 Grape Sugar Content under Quality Attributes Using Normalized Differ-
454 ence Vegetation Index Data and Automated Machine Learning. Sensors
455 22. doi:10.3390/s22093249.
- 456 Kawasaki, K., Uchida, S., 2016. Quality Matters More Than Quan-
457 tity: Asymmetric Temperature Effects on Crop Yield and Quality
458 Grade. American journal of agricultural economics 98, 1195–1209.
459 doi:10.1093/ajae/aaw036.
- 460 Keith Jones, 2002. Australian Wine Industry Environment Strategy.
- 461 Knight, H., Megicks, P., Agarwal, S., Leenders, M., 2019. Firm resources and
462 the development of environmental sustainability among small and medium-
463 sized enterprises: Evidence from the Australian wine industry. Business
464 Strategy and the Environment 28, 25–39. doi:10.1002/bse.2178.
- 465 Kuhn, M., 2008. Building Predictive Models in R Using the
466 caret Package. Journal of Statistical Software, Articles 28, 1–26.
467 doi:10.18637/jss.v028.i05.
- 468 Laurent, C., Le Moguédec, G., Taylor, J., Scholasch, T., Tisseyre, B., Metay,
469 A., 2022. Local influence of climate on grapevine: An analytical pro-
470 cess involving a functional and Bayesian exploration of farm data time
471 series synchronised with an eGDD thermal index. OENO one 56, 301–317.
472 doi:10.20870/oenone.2022.56.2.5443.

- 473 Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D.,
474 2018. Machine Learning in Agriculture: A Review. Sensors 18.
475 doi:10.3390/s18082674.
- 476 Lin, H., 2012. Hydropedology : Synergistic Integration of Soil Science and
477 Hydrology. Elsevier Science & Technology, San Diego, NETHERLANDS,
478 THE.
- 479 Longbottom, M., Petrie, P., 2015. Role of vineyard practices in generating
480 and mitigating greenhouse gas emissions. Australian Journal of Grape and
481 Wine Research 21, 522–536. doi:10.1111/ajgw.12197.
- 482 Mariani, A., Vastola, A., 2015. Sustainable winegrowing: Current perspec-
483 tives. International Journal of Wine Research 7, 37–48.
- 484 Mejean Perrot, N., Tonda, A., Brunetti, I., Guillemin, H., Perret, B.,
485 Goulet, E., Guerin, L., Picque, D., 2022. A decision-support sys-
486 tem to predict grape berry quality and wine potential for a Chenin
487 vineyard. Computers and electronics in agriculture 200, 107167.
488 doi:10.1016/j.compag.2022.107167.
- 489 Montalvo-Falcón, J.V., Sánchez-García, E., Marco-Lajara, B., Martínez-
490 Falcó, J., 2023. Sustainability Research in the Wine Industry: A Bib-
491 liometric Approach. Agronomy 13. doi:10.3390/agronomy13030871.
- 492 Oliver, D., Bramley, R., Riches, D., Porter, I., Edwards, J., 2013. Review:
493 Soil physical and chemical properties as indicators of soil quality in Aus-
494 tralian viticulture. Australian Journal of Grape and Wine Research 19,
495 129–139. doi:10.1111/ajgw.12016.

- 496 Pisciotta, A., Cusimano, G., Favara, R., 2015. Groundwater nitrate risk
497 assessment using intrinsic vulnerability methods: A comparative study
498 of environmental impact by intensive farming in the Mediterranean re-
499 gion of Sicily, Italy. *Journal of geochemical exploration* 156, 89–100.
500 doi:10.1016/j.gexplo.2015.05.002.
- 501 R Core Team, 2021. R: A Language and Environment for Statistical Com-
502 putting. R Foundation for Statistical Computing.
- 503 Santiago-Brown, I., Metcalfe, A., Jerram, C., Collins, C., 2015. Sustain-
504 ability Assessment in Wine-Grape Growing in the New World: Economic,
505 Environmental, and Social Indicators for Agricultural Businesses. *Sustain-
506 ability* 7, 8178–8204. doi:10.3390/su7078178.
- 507 SOAR, C., SADRAS, V., PETRIE, P., 2008. Climate drivers of red wine
508 quality in four contrasting Australian wine regions. *Australian journal of
509 grape and wine research* 14, 78–90. doi:10.1111/j.1755-0238.2008.00011.x.
- 510 Srivastava, S., Sadistap, S., 2018. Non-destructive sensing methods for qual-
511 ity assessment of on-tree fruits: A review. *Journal of Food Measurement
512 and Characterization* 12, 497–526.
- 513 Suarez, L., Zhang, P., Sun, J., Wang, Y., Poblete, T., Hornero, A.,
514 Zarco-Tejada, P., 2021. Assessing wine grape quality parameters
515 using plant traits derived from physical model inversion of hyper-
516 spectral imagery. *Agricultural and forest meteorology* 306, 108445.
517 doi:10.1016/j.agrformet.2021.108445.

- 518 Sustainable Winegrowing Australia, SWA., 2021. Sustainable Winegrowing
519 Australia User Manual.
- 520 SWA, S.W.A., 2022. Sustainable Wingrowing Australia.
521 <https://sustainablewinegrowing.com.au/case-studies/>.
- 522 Wang, Y., He, Y.N., He, L., He, F., Chen, W., Duan, C.Q., Wang,
523 J., 2019. Changes in global aroma profiles of Cabernet Sauvignon in
524 response to cluster thinning. Food research international 122, 56–65.
525 doi:10.1016/j.foodres.2019.03.061.
- 526 Wine Australia, 2019. National Vintage Report 2019 .
- 527 Wine Australia, 2020. National Vintage Report 2020 .
- 528 Wine Australia, 2021. National Vintage Report 2021 .
- 529 Wine Australia, 2022. National Vintage Report 2022 .
- 530 Winemakers' Federation of Australia, 2013. National Vintage Report 2013 .
- 531 Winemakers' Federation of Australia, 2014. National Vintage Report 2014 .
- 532 Winemakers' Federation of Australia, 2015. National Vintage Report 2015 .
- 533 Winemakers' Federation of Australia, 2016. National Vintage Report 2016 .
- 534 Winemakers' Federation of Australia, 2017. National Vintage Report 2017 .
- 535 Winemakers' Federation of Australia, 2018. National Vintage Report 2018 .

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

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

539 ZHU, D.w., ZHANG, H.c., GUO, B.w., XU, K., DAI, Q.g., WEI, H.y., GAO,
 540 H., HU, Y.j., CUI, P.y., HUO, Z.y., 2017. Effects of nitrogen level on yield
 541 and quality of japonica soft super rice. Journal of Integrative Agriculture
 542 16, 1018–1027. doi:10.1016/S2095-3119(16)61577-0.

543 **Appendix A. Appendix**

Model 1 – Normal QQ

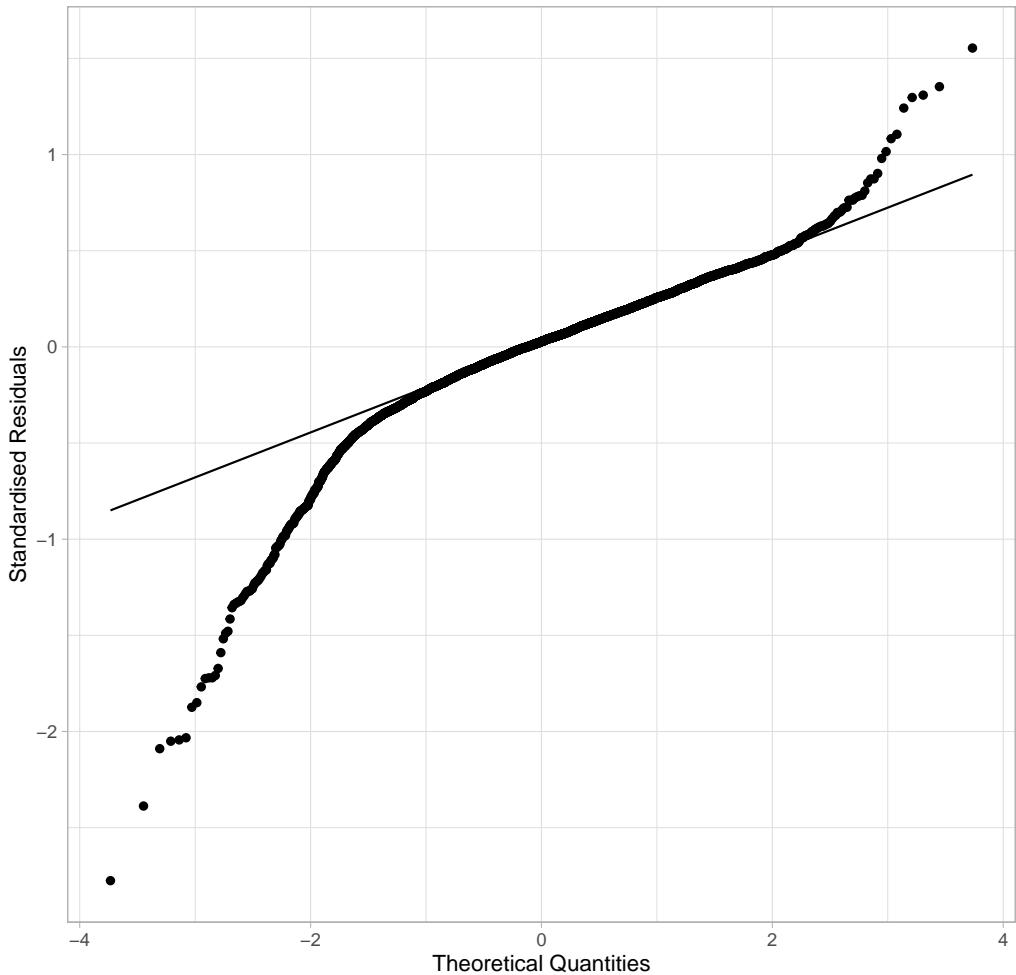


Figure A.5: QQ-plot of Model 1.

Model 1 – Residuals vs fitted

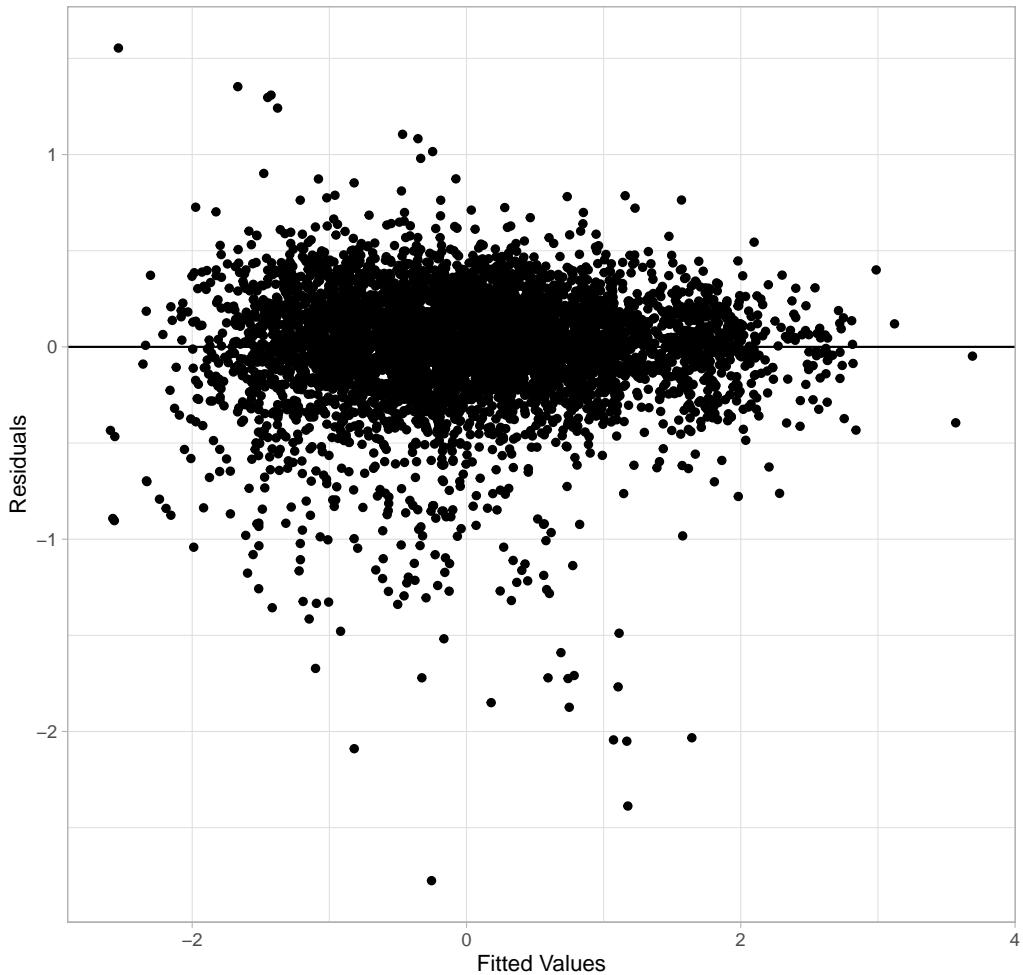


Figure A.6: Residuals vs fitted values for Model 1.

Model 2 – Normal QQ

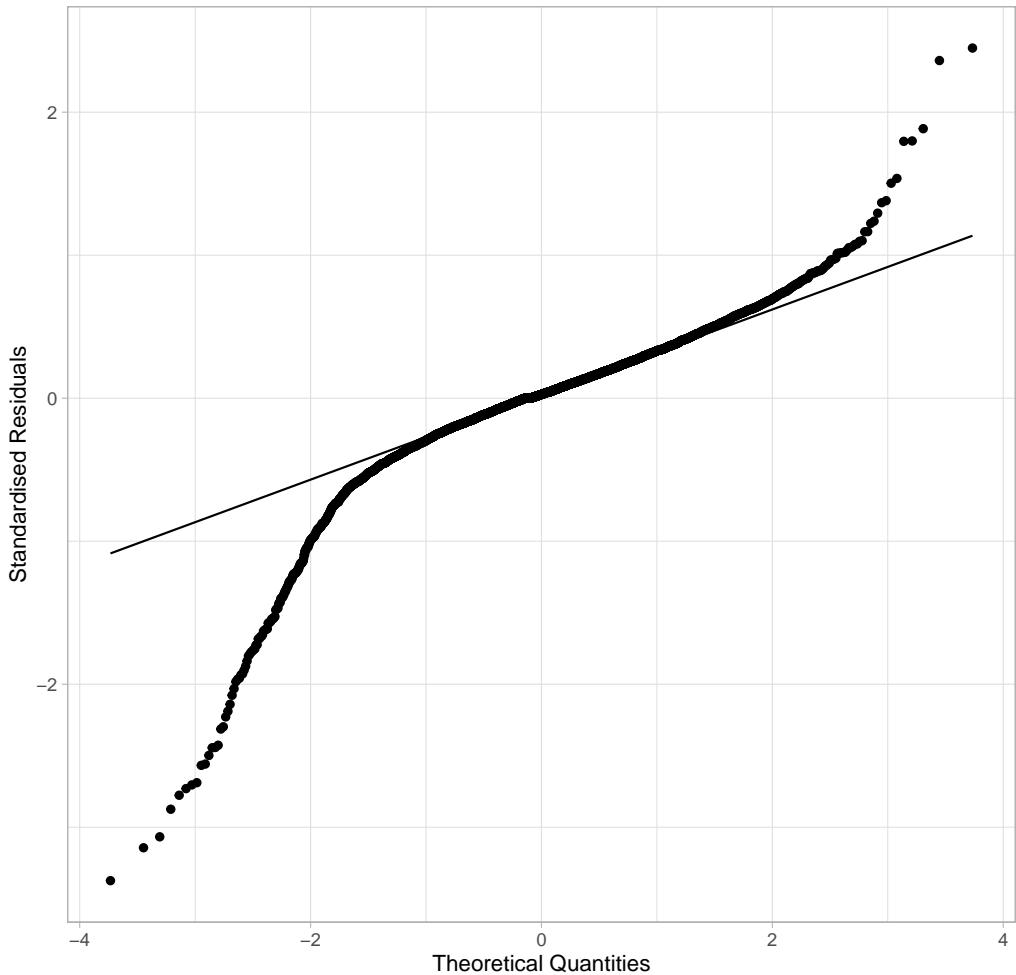


Figure A.7: QQ-plot of Model 2.

Model 2 – Residuals vs fitted

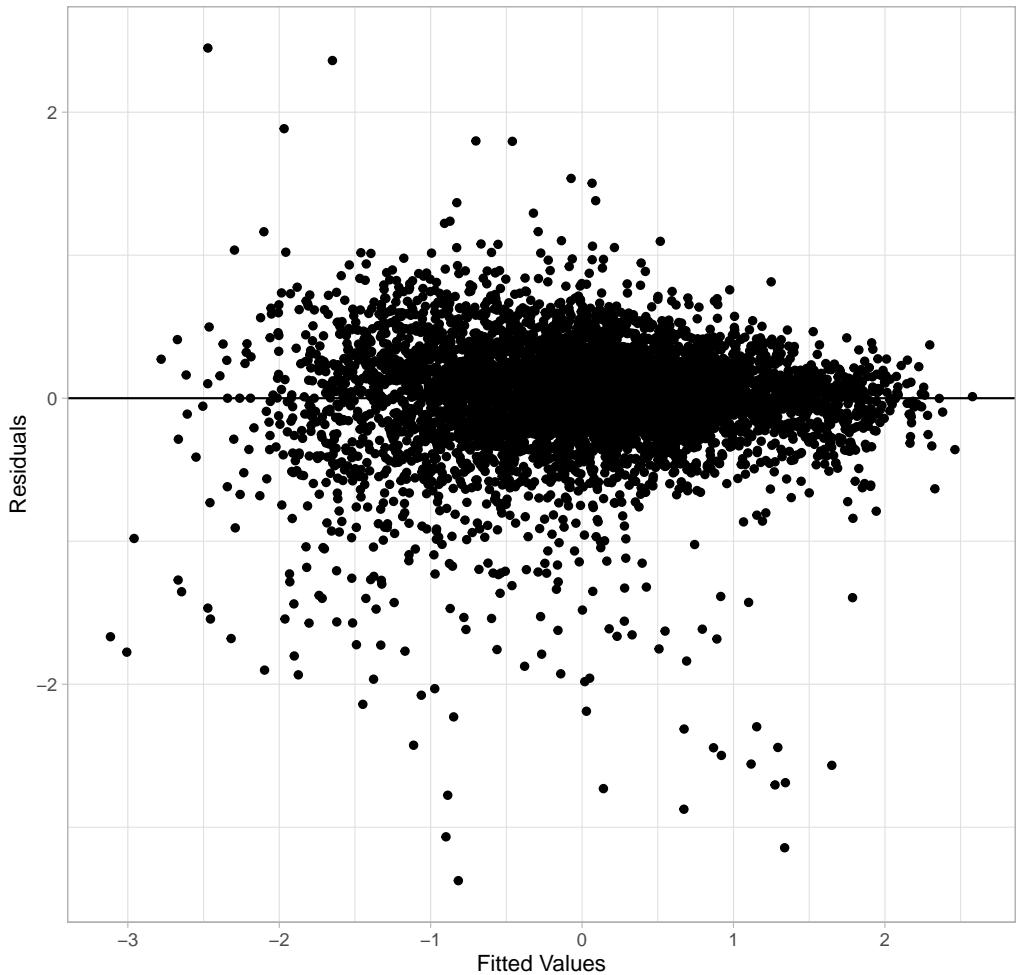


Figure A.8: Residuals vs fitted values for Model 2.

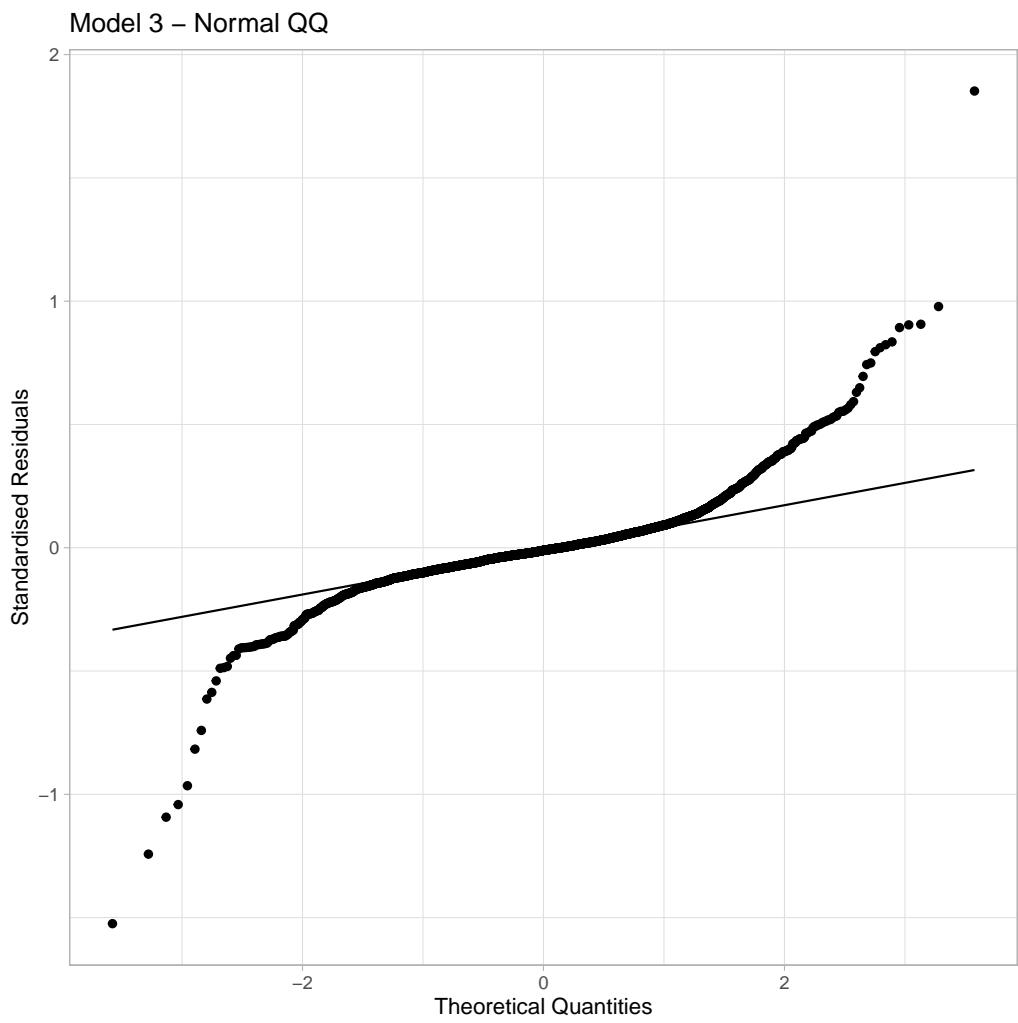


Figure A.9: QQ-plot of Model 3.

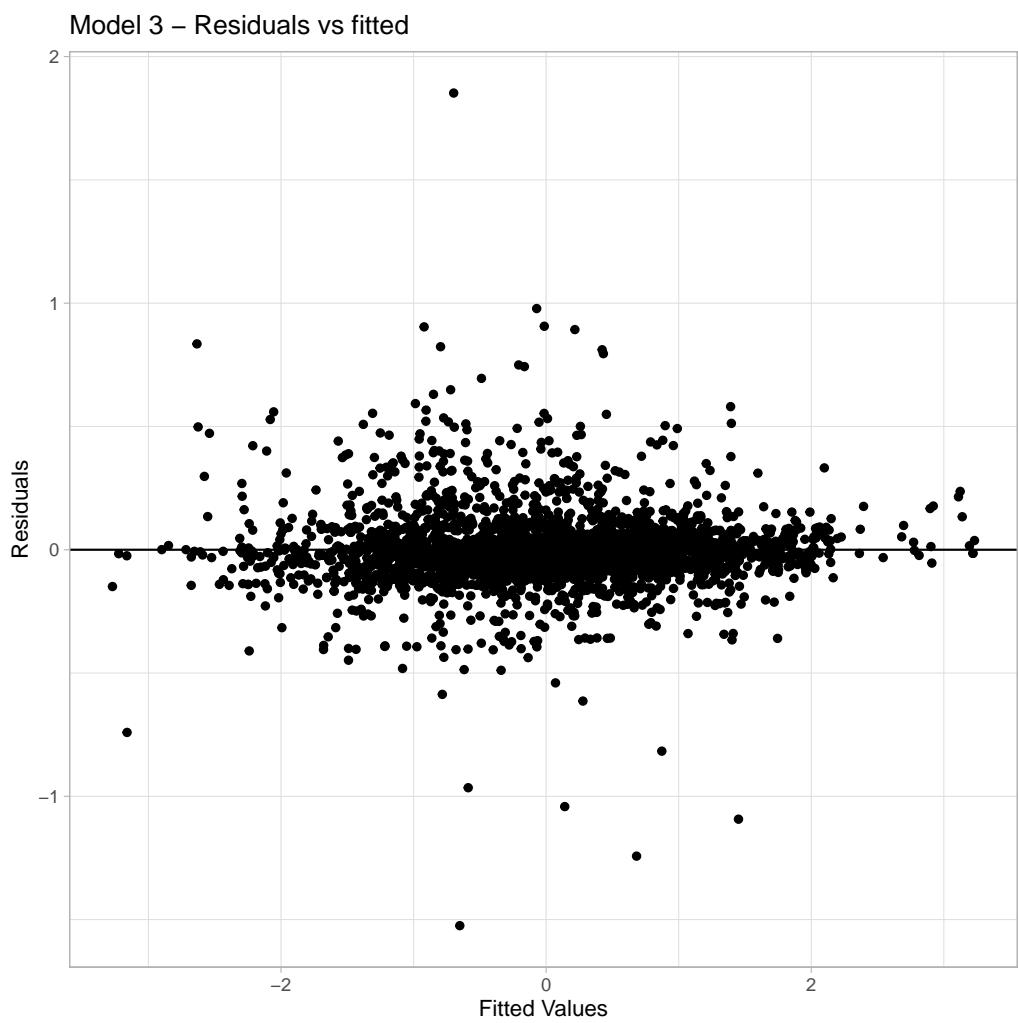


Figure A.10: Residuals vs fitted values for Model 3.

Model 4 – Normal QQ

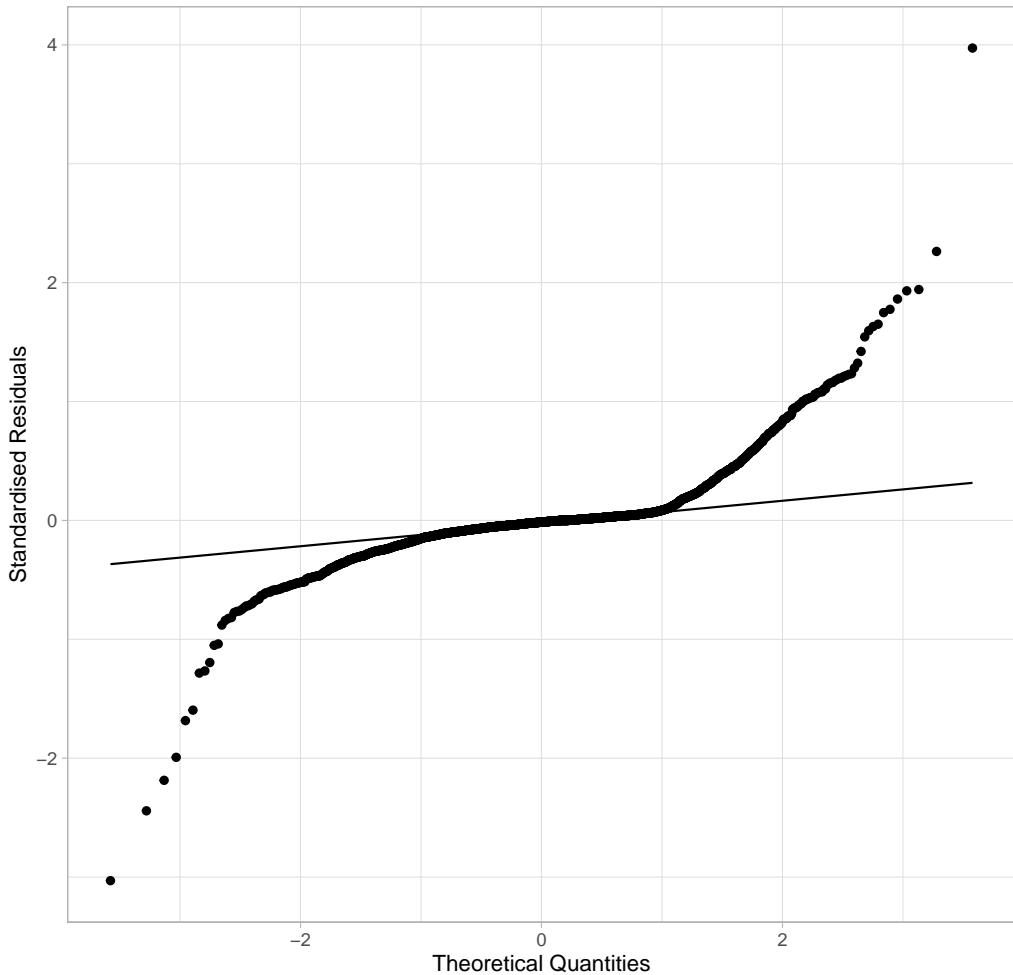


Figure A.11: QQ-plot of Model 4.

Model 4 – Residuals vs fitted

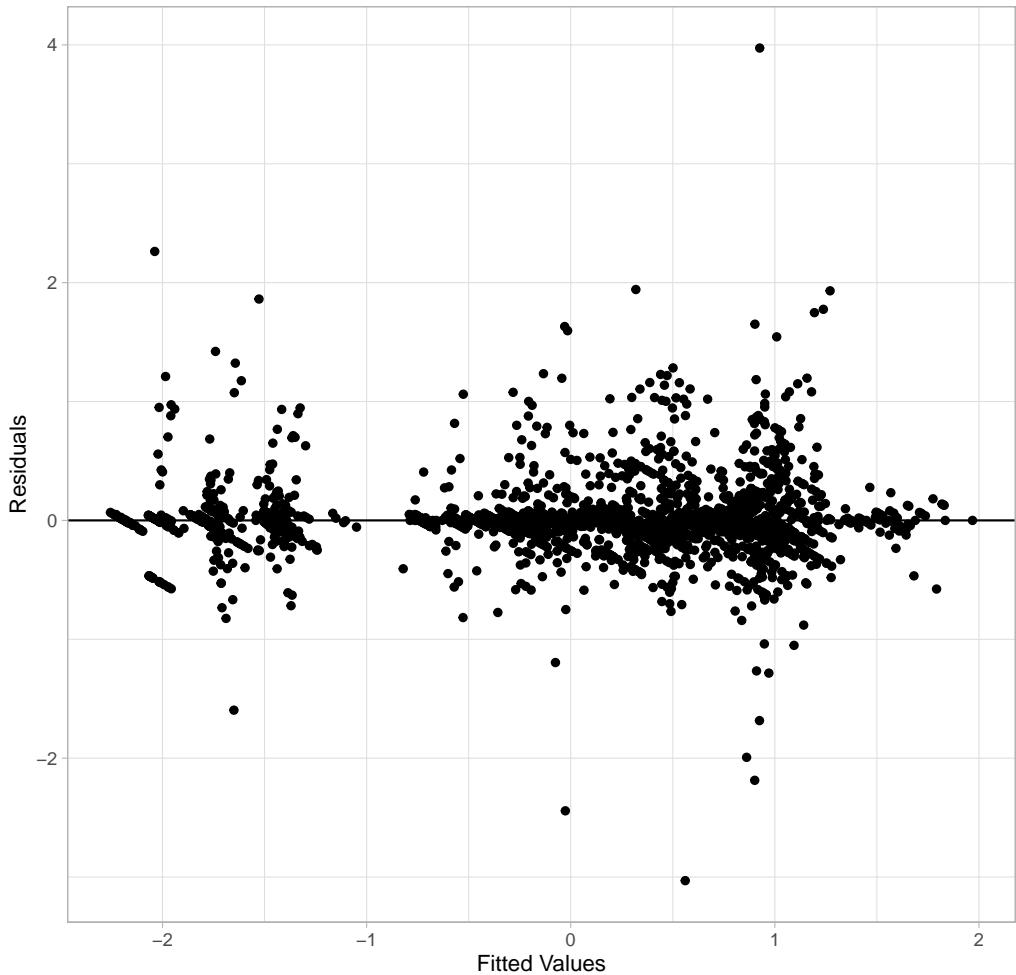


Figure A.12: Residuals vs fitted values for Model 4.

Model 5 – Normal QQ

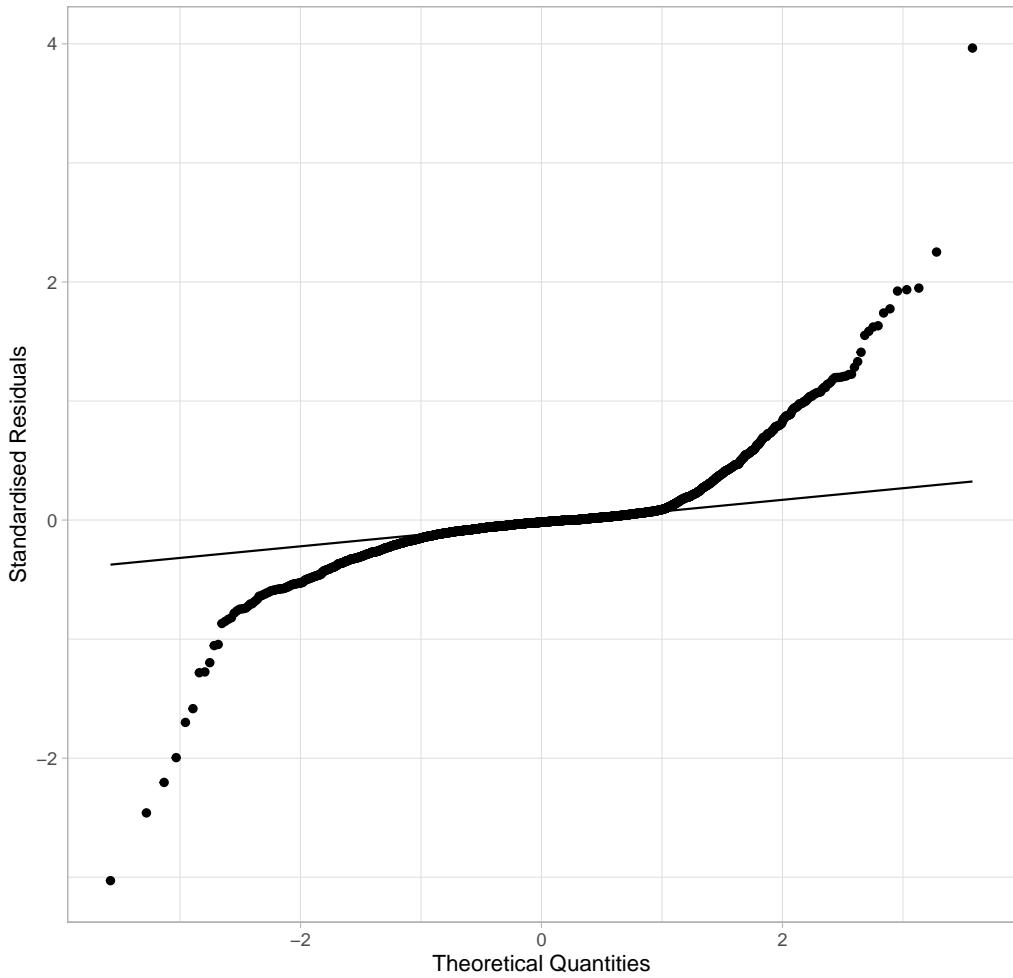


Figure A.13: QQ-plot of Model 5.

Model 5 – Residuals vs fitted

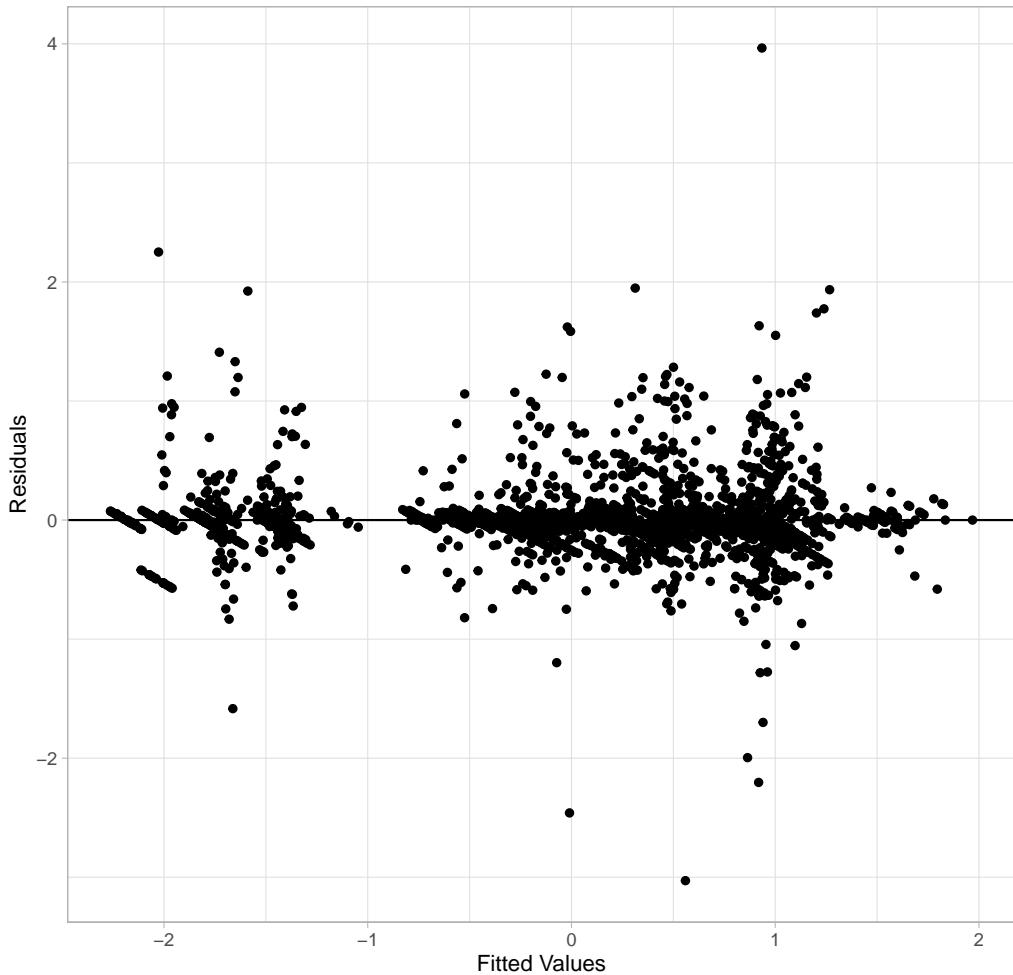


Figure A.14: Residuals vs fitted values for Model 5.