

¹ Highlights

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

⁴ Author

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

¹¹ The influence of resource use on yield versus sale price
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 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.

¹⁵ **1. Introduction**

¹⁶ The global focus on sustainability in agronomic industries has changed
¹⁷ the way in which these enterprises do business. A dilemma exists across
¹⁸ agriculture through shared fundamental considerations of resource use such
¹⁹ as water and fuel, and the resultant yield and crop value that are produced
²⁰ (Hemming et al., 2020; Kawasaki and Uchida, 2016; ZHU et al., 2017). The
²¹ average price of grapes for wine production is driven through its integration
²² within the wine industry. An important connection between grapes and their
²³ price is the grape's perceived quality, which initially defines a wine's potential
²⁴ through the grape's chemical makeup (Black et al., 2015; Schreier and Jen-
²⁵ nings, 1979). Grapes of higher perceived quality or grapes from well known
²⁶ regions are likely to have higher prices (Wine Australia, 2021). Grape quality
²⁷ is connected to the market value of winegrapes, with Wine Australia explic-
²⁸ itly defining grape quality through the use of discrete price brackets in their
²⁹ annual reports (Winemakers' Federation of Australia, 2018). It is important
³⁰ to note that the generalisation made to reflect quality through using aver-
³¹ age price assumes a due diligence of those who purchase the grapes (Yegge,
³² 2001). The economic sustainability of a vineyard is tied to this market cul-
³³ ture, driven by the wine industry. The consideration of sustainability within
³⁴ viticulture is also subject to environmental and sociodemographic pressures
³⁵ (Santiago-Brown et al., 2015). In the Australian context, these pressures
³⁶ include biosecurity and climate and international market changes (Canadell
³⁷ et al., 2021; Longbottom and Petrie, 2015; Oliver et al., 2013).

³⁸ There is an extensive amount of research into the varied effects of factors
³⁹ on grape quality and yield (He et al., 2022; Laurent et al., 2022; Liakos et al.,

40 2018). Due to the lack of long-term and in-depth data there is a lack of
41 research on grape sale price and its driving factors. Furthermore, individual
42 factors are often studied in isolation with yield and sales price not appearing
43 together (Abbal et al., 2016). The lack of consolidated datasets restricts the
44 ability to gain statistical insights at large scales and across multiple regions.
45 As a result, broader studies are lacking (Keith Jones, 2002; Knight et al.,
46 2019). The dataset used for this analysis includes data spanning 10 years
47 from a multitude of vineyards located over a diverse range of Australian
48 winegrowing regions. We use this dataset to describe the relationship of
49 resources related to water and fuel use with the output yield and average sale
50 price of the resultant product, taking into account the size and location of the
51 vineyard. In a practical sense, this aim is a baseline for comparison: given
52 a vineyard within Australia, one could estimate the comparative efficiency
53 with regard to the trade-off between invested resources, yield and average
54 sale price. This is the first time that such a trade-off has been confirmed
55 explicitly across such varying regions, scales and climates in the Australian
56 winegrowing industry.

57 **2. Methods**

58 *2.1. Data*

59 Data used in these analyses were obtained from Sustainable Winegrow-
60 ing Australia and Wine Australia. Sustainable Winegrowing Australia is
61 Australia's national wine industry sustainability program, which aims to fa-
62 cilitate grape growers and winemakers in demonstrating and improving their
63 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	$\text{Yield} \times \text{Average Sale Price}$	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

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

71 Data recorded by Sustainable Winegrowing Australia were entered manu-
72 ally by winegrowers voluntarily using a web based interface. For each model,
73 vineyards were only included if they recorded all the variables used for the
74 corresponding model (see Table 1). Each vineyard had at least recorded
75 region, harvest year, yield and area harvested. Other variables that were
76 used but not present for every single vineyard were average sale price, water
77 used and fuel used (diesel, petrol, biodiesel and LPG). To enable direct com-
78 parisons between fuels, fuel use was converted to tonnes of carbon dioxide
79 equivalent and collectively referenced to as emissions. All variables were con-
80 tinuous except for harvest year and region, which were categorical variables
81 (Table 1).

82 As data from Sustainable Wine Australia were voluntarily given, missing
83 values were improved using regional average prices from the Wine Australia
84 data. Data obtained from Wine Australia were collected via phone surveys
85 and included: total tonnes purchased, average price per tonne and yearly
86 change in price for region and grape varietal; with the data being publicly
87 available.

88 The dataset was split into two distinct subsets that could be used across

89 the different models. The first subset contained all vineyards and was used for
90 two models (Model 1 and Model 2, see Table 1). The second subset contained
91 vineyards which either recorded a value for average price of sale per tonne
92 through Sustainable Winegrowing Australia, or were within a region with an
93 average price of sale recorded by Wine Australia; this subset was used for
94 three further models (Models 3, 4 and 5, see Table 1). These subsets meant
95 that the data would be limited to samples which had recorded values for the
96 response variables (see Table 1), as every sample/vineyard had a recorded
97 value for yield but not average price of sale per tonne.

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

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

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

113 Data preprocessing was conducted prior to analysis using the Python

114 programming language (G. van Rossum, 1995). Preprocessing included the
115 conversion from fuel to scope one emissions and prior calculations for all con-
116 tinuous variables which included logarithmic transformations, centring and
117 scaling by standard deviation. We converted multiple emission sources into
118 scope one emissions using the equation given from the Australian National
119 Greenhouse Accounts Factors (AGDEE, 2021). The calculation conducted
120 using

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

121 where emissions was the product of the quantity of fuel in litres, Q ,
122 a prescribed Energy Content, EC , and emission factors (as given by the
123 Australian National Greenhouse Accounts factors) of scope one, $EF1$, and
124 scope three, $EF3$, to tonnes of Carbon Dioxide Emission equivalent, $tCO2e$
125 (Department of Climate Change, Energy, the Environment and Water, 2022).

126 Differences in vineyard locations were captured through the use of Geo-
127 graphical Indicator Regions defined by Wine Australia (Halliday, 2009; Oliver
128 et al., 2013; SOAR et al., 2008). Although vineyards generally differed in in-
129 come, there were pronounced differences between regions as shown in Figure
130 1. These differences were likely due to the site of a vineyard being important
131 as it predetermines several physical parameters such as climate, geology and
132 soil, making location a widely considered key determinant of grape yield and
133 average sale price (Abbal et al., 2016; Agosta et al., 2012; Fraga et al., 2017).
134 Each GI Region has its own unique mixture of climatic and geophysical
135 properties that describes a unique winegrowing region within Australia and
136 is a protected trademark under the Wine Australia act (Attorney-General's
137 Department, 2010). Both Wine Australia and Sustainable Winegrowing Aus-

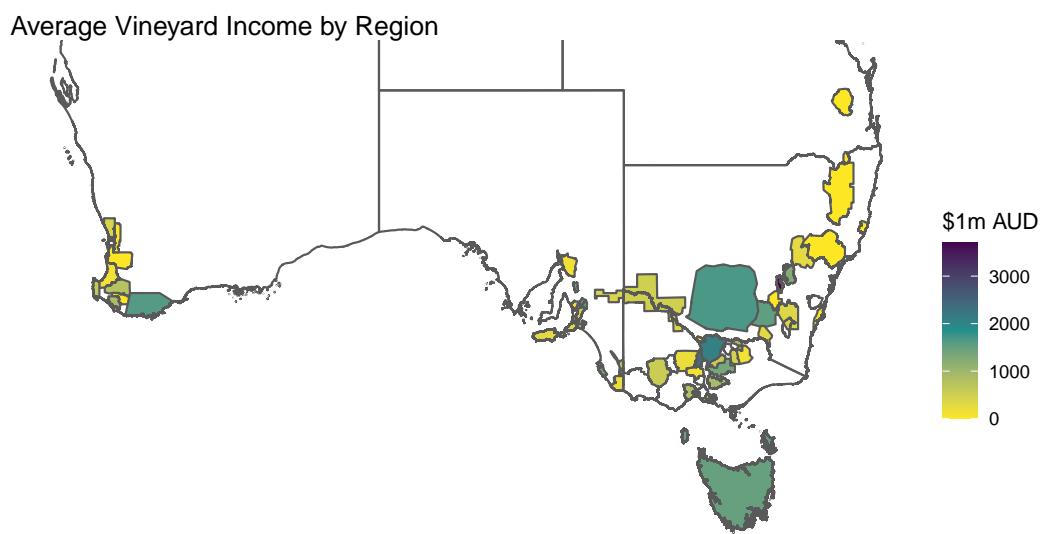


Figure 1: Map of each GI Regions' average income for a vineyard of that region (average grape sale price per tonne \times total tonnes yielded).

138 tralia used the same GI Region categorical variable format to describe loca-
139 tion.

140 The climatic properties of each GI Region were summarised by using pre-
141 defined classifications as per the Sustainable Winegrowing Australia (2021)
142 user manual. The user manual describes climates by rainfall and tempera-
143 ture, creating supersets of Regions of similar climatic properties. The cli-
144 matic groups were used to illustrate similarities and differences occurring in
145 areas larger than GI Regions.

146 *2.2. Analysis*

147 Pairwise Pearson Correlation Coefficients (PPCC) were calculated to as-
148 sess the potential existence of linear relationships between the input and
149 predicted variables. To determine if a coefficient was indicative of a strong
150 relationship, both the magnitude of the value and the associated confidence
151 intervals were evaluated. P-values reflected the significance of a given corre-
152 lation coefficient with statistical significance being declared when the corre-
153 sponding P-value was lower than 0.05. PPCC analyses were undertaken for
154 data on the original scale and for data as a logarithmic transform. Trans-
155 forming data prior to calculating the coefficients changes several things. The
156 logarithmic transform of the data alters the interpretation of the coefficients
157 to percentage change; a coefficient will be indicative of the change in per-
158 centage of one variable compared to the other, scaling by standard deviation
159 also changes this interpretation to be a percentage of that variables standard
160 deviation. When considering the logarithmically transformed variables, a co-
161 efficient of 1 would indicate that the change of one variable by one percentage
162 of its standard deviation would correlate to the other variable changing by

163 one percent of its own standard deviation. The importance of this is the
164 dimensionless nature of these relationships and that it can be translated di-
165 rectly to any vineyard's case that has a well known distribution.

166 Five general linear models (GLMs) were created (see Table 1). GLMs were
167 chosen as they offer the ability to produce statistical models that are explicit
168 in the relationships between predictors and response variables. General Lin-
169 ear Models also allowed the exploration of interactions between predictors
170 and allow for easily comparable differences in the influence and magnitude
171 of relationships. Model fit was measured in terms of R^2 and adjusted R^2
172 as well as F statistics, and t-tests were used to determine if predictors sig-
173 nificantly contributed to the models when accounting for other variables,
174 showing which specific years and areas contributed significantly. Both the
175 PPCC and GLM analyses were created using the R statistical programming
176 language (R Core Team, 2021) with the Caret package (Kuhn, 2008).

177 A variety of alternate methods were also explored, including splines, hi-
178 erarchical regression, general additive models, and generalised linear models.
179 These alternative approaches were not used as final models due to offering
180 little further insights or improvements in accuracy.

181 *2.3. Model Validation*

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

Table 2: Summary statistics of each continuous variable.

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

¹⁸⁷ **3. Results**

¹⁸⁸ *3.1. Exploratory Analysis*

¹⁸⁹ Table 2 shows the summary statistics of each variable in its original units.

¹⁹⁰ The range of these values shows the level of difference between some vine-
¹⁹¹ yards, with operations differing by orders of magnitude in size, yield and
¹⁹² average price of sale (See Table 1).

¹⁹³ 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. All correlations were found to be statistically significant ($p < 2.200E-16$), 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 (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 significance of F-tests ($p < 2.200E-16$). 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

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

209 emissions in Model 3 ($p=0.22$) and Model 4 ($p=0.0.39$), and the interaction
 210 between area harvested and water used in model 2 ($p=0.22$). Note that scope
 211 one emissions was included in all models to directly compare the response
 212 variables as ratios of vineyard size to raw values and because it was strongly
 213 correlated to the response variable in every model (except model 5) especially
 214 for Models 1 and 4 (Table 3).

215 Coefficients related to continuous variables are summarised in Table 5.
 216 In Model 1 all coefficients except for the intercept significantly contributed
 217 to the model ($p < 0.05$), and in Model 2 all coefficients were statistically
 218 significant ($p < 0.05$). However, scope one emissions did not significantly
 219 contribute to models 3, 4 and 5. Model 4 only had statistically significant
 220 contributions from the intercept and water use. Although the coefficient for
 221 water use was statistically significant for each model, it did not have the
 222 highest value; instead, area harvested, being an order of magnitude greater,
 223 dominated the models. Model 5 achieved a similar R^2 to Model 4 without
 224 area harvested, having stronger influences from water use and scope one

Table 5: Summary of each Model's 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		

225 emissions.

226 The regression coefficients for the Year variable of each model is depicted
 227 in Figure 2. For each model, the first year for a model's was used as the
 228 baseline: 2012 for Models 1 and 2, and 2015 for Models 3, 4 and 5. Adelaide
 229 Hills is used as the regional baseline with the interaction between year and
 230 region using the first year and the Adelaide Hills as the baseline. Region and
 231 year contributed, in some but not all cases, more than the other variables.
 232 However, the coefficients varied substantially over the years, and were not
 233 significantly different from zero in some years. Models 4 and 5 are very similar
 234 in Figure 2, indicating that the exclusion of area does not greatly affect the
 235 contribution from yearly influence. Models 4 and 5 have the most prominent
 236 trends, showing an increase in yearly effects over time, with Model 3 also
 237 increasing from 2016 to 2018, but plateauing afterwards. Model 1 and 2 do

238 not show a clear trend, but do drop during 2017 and 2018 after increasing in
239 the first 3 years.

240 Regional differences by temperature and rainfall are summarised in Fig-
241 ure 3. The most notable difference is between vineyards within 'Hot' and
242 'Very Dry' regions (warm inland regions), where high average sale prices are
243 historically low, and yield is high. Water Use changes dramatically between
244 these regions as well, with water being a driving force for yield but not nec-
245 essarily average sale price. The warmer and drier regions tend to also have
246 larger vineyards. These regional differences are further shown in Figure 4, as
247 a ratio of vineyard; where there is a pronounced difference in 'Hot' and 'Very
248 Dry' regions producing more per area with lower average sale prices per area
249 than other regions.

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

254 Table 3.2 shows the validation results of each of the models. The R^2
255 measures of fit show similar results to the initial models, with a slight decrease
256 as expected, indicating that the models are robust and consistent.

257 4. Discussion

258 There was an expected strong relationship between size and resource use,
259 with the overall area of a vineyard and its access to resources greatly deter-
260 mining the upper limit of potential yield. However, size was also inversely
261 related to the potential average sale price. Higher average sales prices were

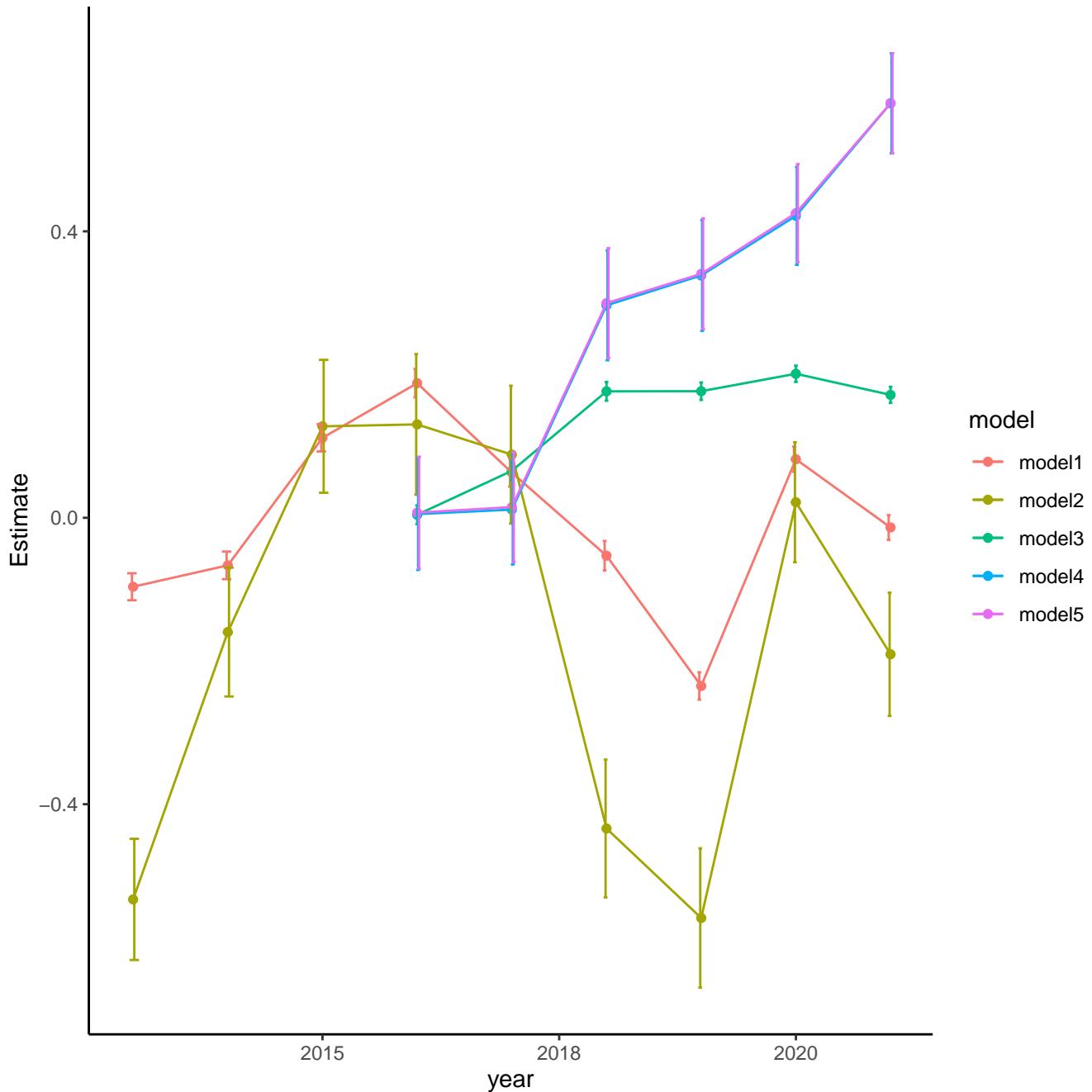


Figure 2: Model Coefficient values for Year, with standard error bars.

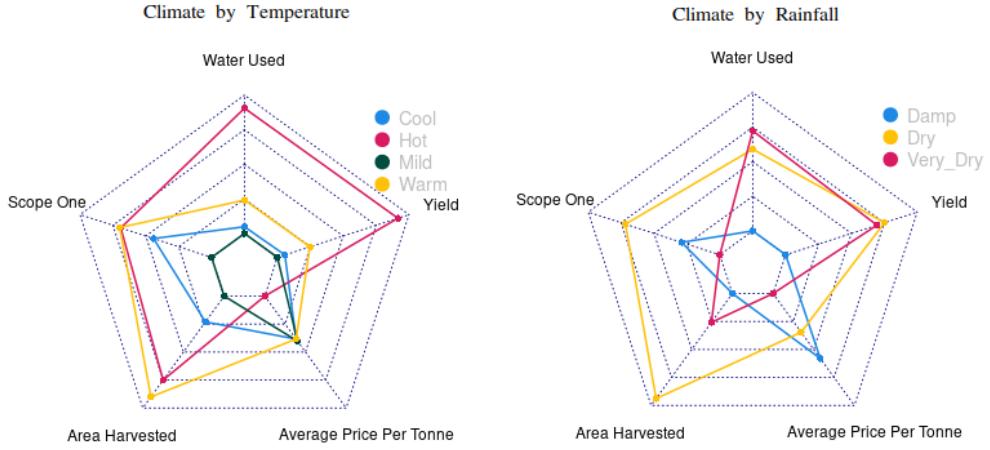


Figure 3: Radar plot of climatic profile's resource use, yield and average sale price. The left reflects vineyards in different climatic temperatures. The right reflects vineyards in different rainfall climates.

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

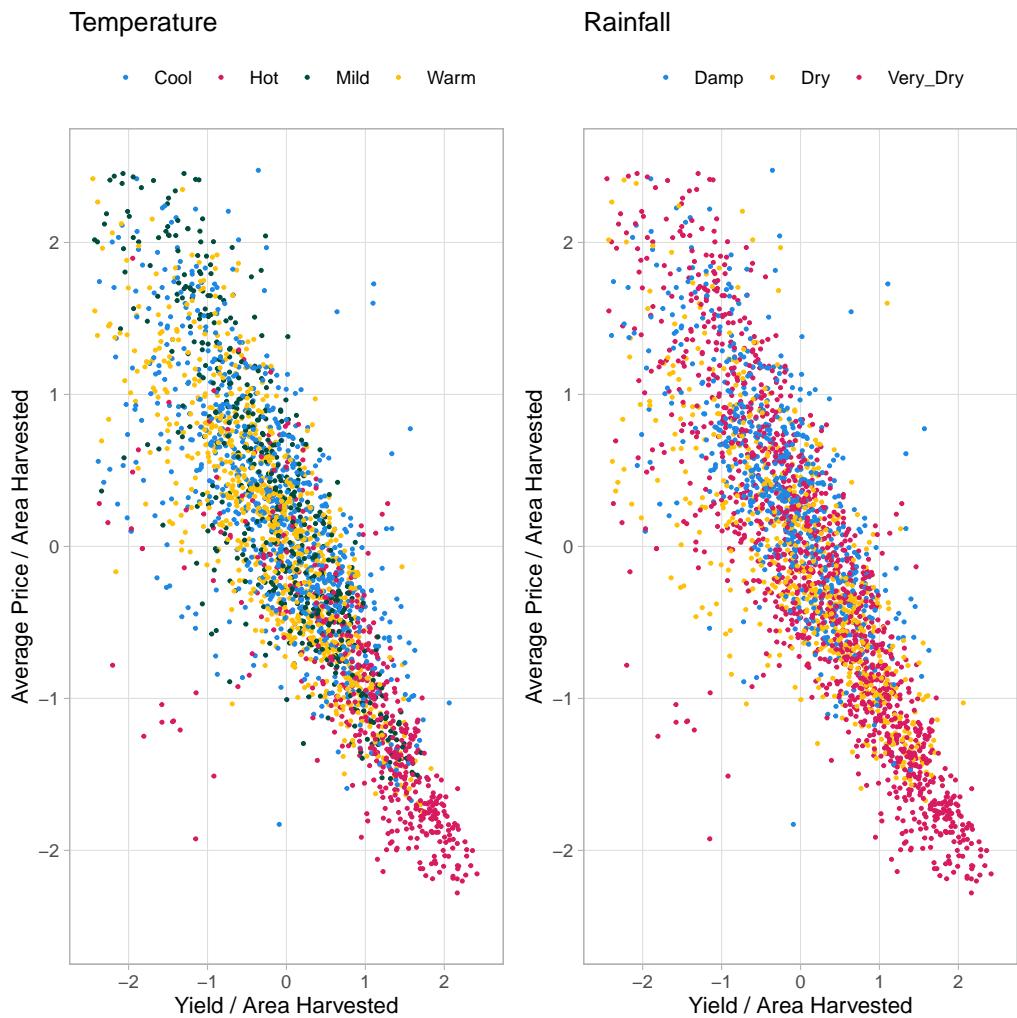


Figure 4: Scatter plot of vineyard yield against the average sale price as ratios to area harvested. The axes are in standard deviations with points coloured by climate.

262 also related to high resource inputs per area instead of the overall expenditure
263 of resources. Vineyard yields and sales price changed greatly by region
264 and year. Even given regional and yearly changes, there was a strong connection
265 between smaller vineyards and higher sales prices. This could have
266 been due to more attention available when managing smaller properties.

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

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

287 motivating the call for more granular data and more sophisticated modelling.

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

306 Climatic pressures are an important consideration for growers, especially
307 those in warmer and drier regions. The Wine Australia reports also show
308 that warm inland regions have seen a decline in profit over the past decade,
309 whereas regions with lower average sales prices did not (Wine Australia, 2019,
310 2020, 2021; Winemakers' Federation of Australia, 2013, 2014, 2015, 2016,
311 2017, 2018). Vineyards in warm inland regions also tend to contain larger

312 vineyards, making up for lower sale prices with larger yields. Considering
313 the negative correlation of average price to area, for this strategy to work,
314 economies of scale become and important factor. Given the large quantities
315 of grapes that can be produced by some vineyards, even at low margins there
316 is the potential to be profitable. However, the increasing climatic pressures
317 mixed with the requirement for larger volumes of water, make the sustain-
318 ability of some vineyards come into question. Furthermore, intensive farming
319 in general is known to jeopardise the sustainability of an operation through
320 the degradation of soil and waterways (Capello et al., 2019; Lin, 2012; Pis-
321 ciotta et al., 2015). There are established methods that can help to mitigate
322 these effect, such as the use of cover crops and midrow crop rotation. How-
323 ever, it has become more apparent that the active reduction of grape yield,
324 through methods such as thinning, can help improve soil health (Condurso
325 et al., 2016; Wang et al., 2019).

326 Some regions appeared to produce grapes of lower average sale price at
327 scale whilst others focussed on producing higher priced grapes in lower vol-
328 umes. This empirical finding is consistent with Wine Australia's annual
329 reports, which shows that some GI regions, such as the Riverland, are known
330 for producing large amounts of lower grade (low value per tonne) grapes
331 (Wine Australia, 2022; Winemakers' Federation of Australia, 2017). Com-
332 paratively, other regions, such as Tasmania, only produce grapes of higher
333 sales price but in smaller quantities. The difference in pricing per tonne
334 between the lowest and highest regional average sales prices was almost a
335 hundred times, showing that region had a profound influence. Some regions
336 also had a mixture of high and low average sales price and yield such as the

337 Yarra Valley, showing that some regions were not solely locked into choosing
338 a focus on high yields or high average sales price. A further possibility is
339 the existence of regional upper limits on potential sales price, or that there
340 are diminishing returns in some regions when pursuing higher sales prices or
341 quantity; however these types of relationships may be obfuscated by knowl-
342 edgeable winegrowers who avoid such pitfalls.

343 Due to regional differences, different strategies are employed, such as
344 some regions targeting mass production over higher sales price. This is most
345 notable when grouping regions by climate, especially when considering GI
346 Regions in the 'Hot Very Dry' climate (see Figure 4). Figure 3 also shows that
347 comparatively 'Warm' and 'Dry' Regions manage their resources incredibly
348 efficiently having generally larger areas using similar resources to those in
349 'Cool' regions but having areas comparable to regions in 'Hot' climates. The
350 coefficients for Model 4 also show a greater benefit of resource use per area
351 when producing grapes at higher average sale prices, showing higher resource
352 use per area reflective of higher average sales price (see Table 2). Although
353 not chosen over GI region, climate was considered to be a large determinant
354 of the ability to produce larger quantities of grapes, as well as a determinant
355 in grape sale price (Agosta et al., 2012). The more granular GI Region likely
356 explained a broader mix of geographical phenomenon, such as soil, geology
357 and access to water resources (Abbal et al., 2016; Carmona et al., 2011). The
358 interaction between year and GI Region likely accounted for events such as
359 bushfires, which would be impactful, but only at a local level, both in time
360 and space.

361 We identified two main limitations to our linear modelling. First, model

362 1 and 2 over-predicting yield may have been due to preventative measures
363 brought on by regional pressures such as fire, frost and disease. More fuel
364 and water was likely used to prevent these issues from spreading within a
365 region, thus disproportionately affecting some vineyards compared to oth-
366 ers locally. This type of maintenance is not well captured in the models,
367 especially when considering that some regions, especially those in warmer
368 areas, are not as prone to disease as cooler climates and could potentially
369 have lower fuel and water use per hectare. This could create a discrepancy in
370 vineyards that utilised preventative measures in wetter regions, as opposed
371 to those that did not, thus expending less fuel and energy but risking disease.
372 When reviewing the differences between regions, it is important to consider
373 that vineyards in 'Hot Very Dry' areas can be hundreds of times the size of
374 those in other regions. This limitation could be overcome by incorporating
375 the profitability of vineyards, comparing the financial success of working at
376 different operational scales.

377 The second limitations was the lack of further explanatory variables to
378 help link models to causal affects. Variables such as the utilisation of renew-
379 able energy, contractors, and the occurrence of disease, fire and frost were
380 originally explored to capture the discrepancies between similar vineyards
381 that produced different yields and crop values. However, none of these vari-
382 ables was significantly correlated with the response variables, and did not
383 add to model accuracy, even when considered as interactions. Allowance
384 for nonlinear relationships, specifically through splines, resulted in more nor-
385 mally distributed residuals but at a drastically reduced overall accuracy when
386 comparing R^2 and Residual Square Error. Attempts to fully explain small

387 variations was always overshadowed by the dramatic differences in regional
388 trends. Having more data for each region would also be beneficial, allowing
389 greater comparison between regions. More variables, such as soil quality and
390 health, may also help to discern vineyards that can produce larger volumes
391 of grapes at higher prices.

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

400 5. Conclusion

401 This study delved into the relationships between resource use, grape sales
402 price and yield. The findings underscore the multifaceted nature of vineyard
403 management, where the interplay of size, resource allocation, climate, and
404 regional influences collectively shape both the expected sale price and the
405 quantity of grape yields. Average sales price of grapes was not solely tied
406 to the overall expenditure of resources, but rather to the efficient allocation
407 of resources per area. This emphasises that factors beyond sheer scale con-
408 tribute significantly to the final sale price of grapes produced. Moreover, re-
409 gional and yearly variations exhibited substantial effects on vineyard outputs,
410 impacting both sales price and quantity. The connection observed between

⁴¹¹ smaller vineyards and higher grape sale price suggests that the management
⁴¹² of smaller properties might be more streamlined and effective.

⁴¹³ **References**

⁴¹⁴ Abad, J., Hermoso de Mendoza, I., Marín, D., Orcaray, L., Santesteban, L.G., 2021. Cover crops in viticulture. A systematic review (1):
⁴¹⁶
Implications on soil characteristics and biodiversity in vineyard.
⁴¹⁷ OENO One 55, 295–312. doi:10.20870/oenone.2021.55.1.3599.

⁴¹⁸ Abbal, P., Sablayrolles, J.M., Matzner-Lober, É., Boursiquot, J.M., Baudrit, C., Carboneau, 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.

⁴²² AGDEE, A.G.D.o.t.E.a.E., 2021. National Greenhouse Accounts Factors:
⁴²³ 2021.

⁴²⁴ Agosta, E., Canziani, P., Cavagnaro, M., 2012. Regional climate variability
⁴²⁵ impacts on the annual grape yield in Mendoza, Argentina. Journal of Applied Meteorology and Climatology 51, 993–1009.

⁴²⁷ Attorney-General's Department, 2010. Wine Australia Corporation Act
⁴²⁸ 1980.

⁴²⁹ Baiano, A., 2021. An Overview on Sustainability in the Wine Production
⁴³⁰ Chain. Beverages 7. doi:10.3390/beverages7010015.

- 431 Black, C., Parker, M., Siebert, T., Capone, D., Francis, I., 2015. Terpenoids
432 and their role in wine flavour: Recent advances. *Australian Journal of*
433 *Grape and Wine Research* 21, 582–600. doi:10.1111/ajgw.12186.
- 434 Canadell, J.G., Meyer, C.P.M., Cook, G.D., Dowdy, A., Briggs, P.R.,
435 Knauer, J., Pepler, A., Haverd, V., 2021. Multi-decadal increase of forest
436 burned area in Australia is linked to climate change. *Nature Communications*
437 12, 6921. doi:10.1038/s41467-021-27225-4.
- 438 Capello, G., Biddoccu, M., Ferraris, S., Cavallo, E., 2019. Effects of Tractor
439 Passes on Hydrological and Soil Erosion Processes in Tilled and Grassed
440 Vineyards. *Water* 11. doi:10.3390/w11102118.
- 441 Carmona, G., Varela-Ortega, C., Bromley, J., 2011. The Use of Participatory
442 Object-Oriented Bayesian Networks and Agro-Economic Models for
443 Groundwater Management in Spain. *Water resources management* 25,
444 1509–1524. doi:10.1007/s11269-010-9757-y.
- 445 Condurso, C., Cincotta, F., Tripodi, G., Sparacio, A., Giglio, D.M.L., Sparla,
446 S., Verzera, A., 2016. Effects of cluster thinning on wine quality of Syrah
447 cultivar (*Vitis vinifera* L.). *European food research & technology* 242,
448 1719–1726. doi:10.1007/s00217-016-2671-7.
- 449 Cortez, P., Teixeira, J., Cerdeira, A., Almeida, F., Matos, T., Reis, J., 2009.
450 Using data mining for wine quality assessment, in: *Discovery Science: 12th*
451 *International Conference, DS 2009, Porto, Portugal, October 3-5, 2009* 12,
452 Springer. pp. 66–79.

- 453 Department of Climate Change, Energy, the Environment and Water, 2022.
- 454 Australian National Greenhouse Accounts Factors.
- 455 Fraga, H., Costa, R., Santos, J.A., 2017. Multivariate clustering of viticultural
456 terroirs in the Douro winemaking region. Ciênc Téc. Vitiv. 32,
457 142–153.
- 458 G. van Rossum, 1995. Python tutorial, Technical Report CS-R9526. Centrum
459 voor Wiskunde en Informatica (CWI),.
- 460 Hall, A., Lamb, D.W., Holzapfel, B.P., Louis, J.P., 2011. Within-season
461 temporal variation in correlations between vineyard canopy and winegrape
462 composition and yield. Precision Agriculture 12, 103–117.
- 463 Halliday, J.C.J.C., 2009. Australian Wine Encyclopedia. Hardie Grant
464 Books, VIC.
- 465 He, L., Fang, W., Zhao, G., Wu, Z., Fu, L., Li, R., Majeed, Y.,
466 Dhupia, J., 2022. Fruit yield prediction and estimation in orchards:
467 A state-of-the-art comprehensive review for both direct and indirect
468 methods. Computers and Electronics in Agriculture 195, 106812.
469 doi:10.1016/j.compag.2022.106812.
- 470 Hemming, S., de Zwart, F., Elings, A., Petropoulou, A., Righini, I., 2020.
471 Cherry tomato production in intelligent greenhouses-sensors and ai for con-
472 trol of climate, irrigation, crop yield, and quality. Sensors (Basel, Switzer-
473 land) 20, 1–30. doi:10.3390/s20226430.
- 474 I. Goodwin,, L. McClymont,, D. Lanyon, A. Zerihun, J. Hornbuckle, M.

- 475 Gibberd, D. Mowat, D. Smith, M. Barnes, R. Correll, 2009. Managing soil
476 and water to target quality and reduce environmental impact.
- 477 Kasimati, A., Espejo-García, B., Darra, N., Fountas, S., 2022. Predicting
478 Grape Sugar Content under Quality Attributes Using Normalized Differ-
479 ence Vegetation Index Data and Automated Machine Learning. Sensors
480 22. doi:10.3390/s22093249.
- 481 Kawasaki, K., Uchida, S., 2016. Quality Matters More Than Quan-
482 tity: Asymmetric Temperature Effects on Crop Yield and Quality
483 Grade. American journal of agricultural economics 98, 1195–1209.
484 doi:10.1093/ajae/aaw036.
- 485 Keith Jones, 2002. Australian Wine Industry Environment Strategy.
- 486 Knight, H., Megicks, P., Agarwal, S., Leenders, M., 2019. Firm resources and
487 the development of environmental sustainability among small and medium-
488 sized enterprises: Evidence from the Australian wine industry. Business
489 Strategy and the Environment 28, 25–39. doi:10.1002/bse.2178.
- 490 Kuhn, M., 2008. Building Predictive Models in R Using the
491 caret Package. Journal of Statistical Software, Articles 28, 1–26.
492 doi:10.18637/jss.v028.i05.
- 493 Laurent, C., Le Moguédec, G., Taylor, J., Scholasch, T., Tisseyre, B., Metay,
494 A., 2022. Local influence of climate on grapevine: An analytical pro-
495 cess involving a functional and Bayesian exploration of farm data time
496 series synchronised with an eGDD thermal index. OENO one 56, 301–317.
497 doi:10.20870/oenone.2022.56.2.5443.

- 498 Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D.,
499 2018. Machine Learning in Agriculture: A Review. Sensors 18.
500 doi:10.3390/s18082674.
- 501 Lin, H., 2012. Hydropedology : Synergistic Integration of Soil Science and
502 Hydrology. Elsevier Science & Technology, San Diego, NETHERLANDS,
503 THE.
- 504 Longbottom, M., Petrie, P., 2015. Role of vineyard practices in generating
505 and mitigating greenhouse gas emissions. Australian Journal of Grape and
506 Wine Research 21, 522–536. doi:10.1111/ajgw.12197.
- 507 Mariani, A., Vastola, A., 2015. Sustainable winegrowing: Current perspec-
508 tives. International Journal of Wine Research 7, 37–48.
- 509 Montalvo-Falcón, J.V., Sánchez-García, E., Marco-Lajara, B., Martínez-
510 Falcó, J., 2023. Sustainability Research in the Wine Industry: A Bib-
511 liometric Approach. Agronomy 13. doi:10.3390/agronomy13030871.
- 512 Oliver, D., Bramley, R., Riches, D., Porter, I., Edwards, J., 2013. Review:
513 Soil physical and chemical properties as indicators of soil quality in Aus-
514 tralian viticulture. Australian Journal of Grape and Wine Research 19,
515 129–139. doi:10.1111/ajgw.12016.
- 516 Pisciotta, A., Cusimano, G., Favara, R., 2015. Groundwater nitrate risk
517 assessment using intrinsic vulnerability methods: A comparative study
518 of environmental impact by intensive farming in the Mediterranean re-
519 gion of Sicily, Italy. Journal of geochemical exploration 156, 89–100.
520 doi:10.1016/j.gexplo.2015.05.002.

- 521 R Core Team, 2021. R: A Language and Environment for Statistical Com-
522 putting. R Foundation for Statistical Computing.
- 523 Santiago-Brown, I., Metcalfe, A., Jerram, C., Collins, C., 2015. Sustain-
524 ability Assessment in Wine-Grape Growing in the New World: Economic,
525 Environmental, and Social Indicators for Agricultural Businesses. Sustain-
526 ability 7, 8178–8204. doi:10.3390/su7078178.
- 527 Schreier, P., Jennings, W.G., 1979. Flavor composition of wines: A re-
528 view. C R C Critical Reviews in Food Science and Nutrition 12, 59–111.
529 doi:10.1080/10408397909527273.
- 530 SOAR, C., SADRAS, V., PETRIE, P., 2008. Climate drivers of red wine
531 quality in four contrasting Australian wine regions. Australian journal of
532 grape and wine research 14, 78–90. doi:10.1111/j.1755-0238.2008.00011.x.
- 533 Srivastava, S., Sadistap, S., 2018. Non-destructive sensing methods for qual-
534 ity assessment of on-tree fruits: A review. Journal of Food Measurement
535 and Characterization 12, 497–526.
- 536 Sustainable Winegrowing Australia, SWA., 2021. Sustainable Winegrowing
537 Australia User Manual.
- 538 SWA, S.W.A., 2022. Sustainable Wingrowing Australia.
539 <https://sustainablewinegrowing.com.au/case-studies/>.
- 540 Wang, Y., He, Y.N., He, L., He, F., Chen, W., Duan, C.Q., Wang,
541 J., 2019. Changes in global aroma profiles of Cabernet Sauvignon in
542 response to cluster thinning. Food research international 122, 56–65.
543 doi:10.1016/j.foodres.2019.03.061.

- 544 Wine Australia, 2019. National Vintage Report 2019 .
- 545 Wine Australia, 2020. National Vintage Report 2020 .
- 546 Wine Australia, 2021. National Vintage Report 2021 .
- 547 Wine Australia, 2022. National Vintage Report 2022 .
- 548 Winemakers' Federation of Australia, 2013. National Vintage Report 2013 .
- 549 Winemakers' Federation of Australia, 2014. National Vintage Report 2014 .
- 550 Winemakers' Federation of Australia, 2015. National Vintage Report 2015 .
- 551 Winemakers' Federation of Australia, 2016. National Vintage Report 2016 .
- 552 Winemakers' Federation of Australia, 2017. National Vintage Report 2017 .
- 553 Winemakers' Federation of Australia, 2018. National Vintage Report 2018 .
- 554 Yegge, J.M., 2001. Influence of Sensory and Non-Sensory Attributes of
555 Chardonnay Wine on Acceptance and Purchase Intent. Ph.D. thesis. Pro-
556 Quest Dissertations Publishing.
- 557 ZHU, D.w., ZHANG, H.c., GUO, B.w., XU, K., DAI, Q.g., WEI, H.y., GAO,
558 H., HU, Y.j., CUI, P.y., HUO, Z.y., 2017. Effects of nitrogen level on yield
559 and quality of japonica soft super rice. Journal of Integrative Agriculture
560 16, 1018–1027. doi:10.1016/S2095-3119(16)61577-0.

561 **Appendix A. Appendix**

Model 1 – Normal QQ

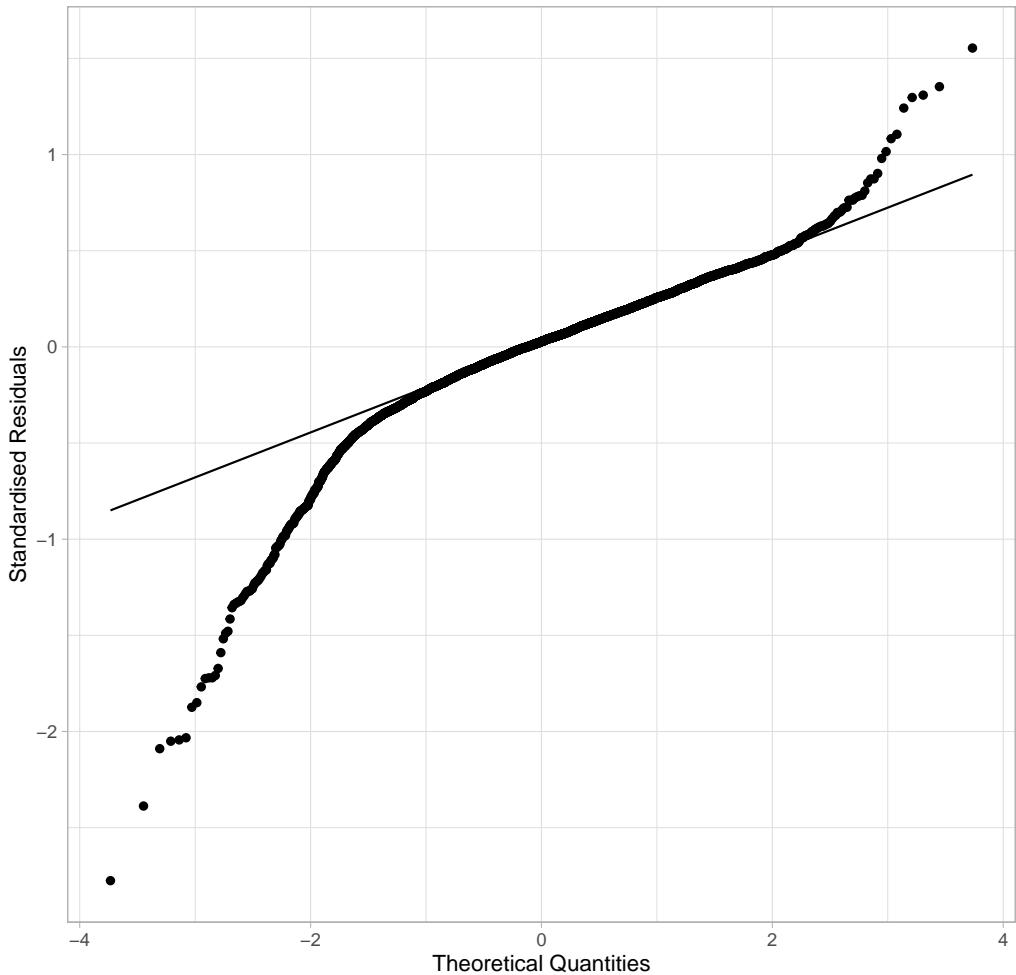


Figure A.5: QQ-plot of Model 1.

Model 1 – Residuals vs fitted

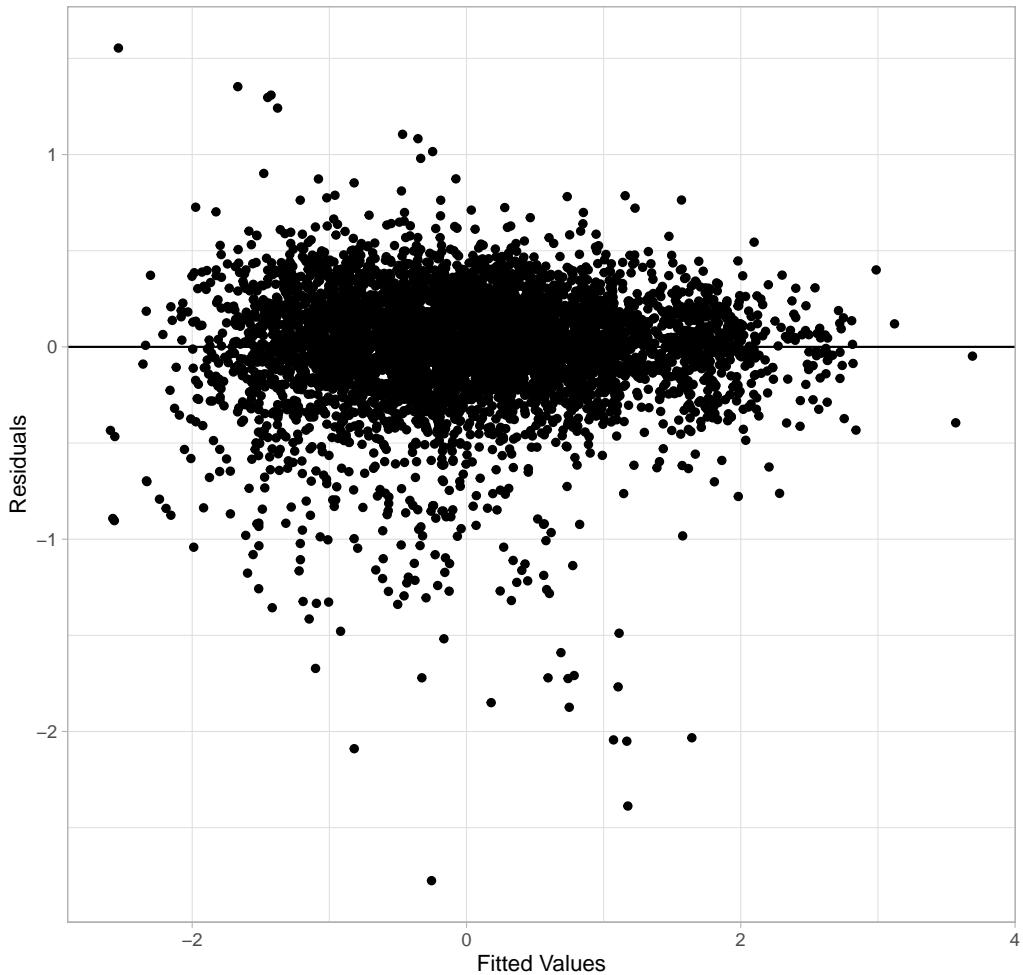


Figure A.6: Residuals vs fitted values for Model 1.

Model 2 – Normal QQ

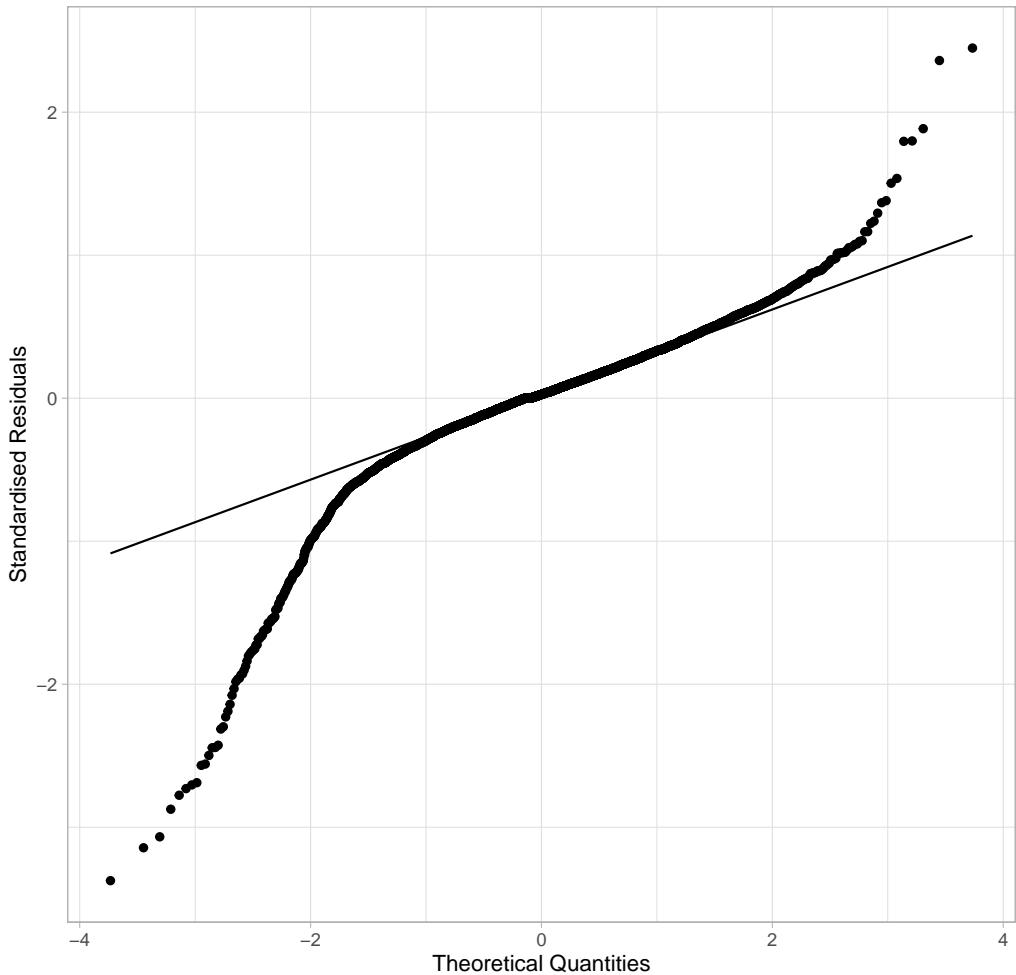


Figure A.7: QQ-plot of Model 2.

Model 2 – Residuals vs fitted

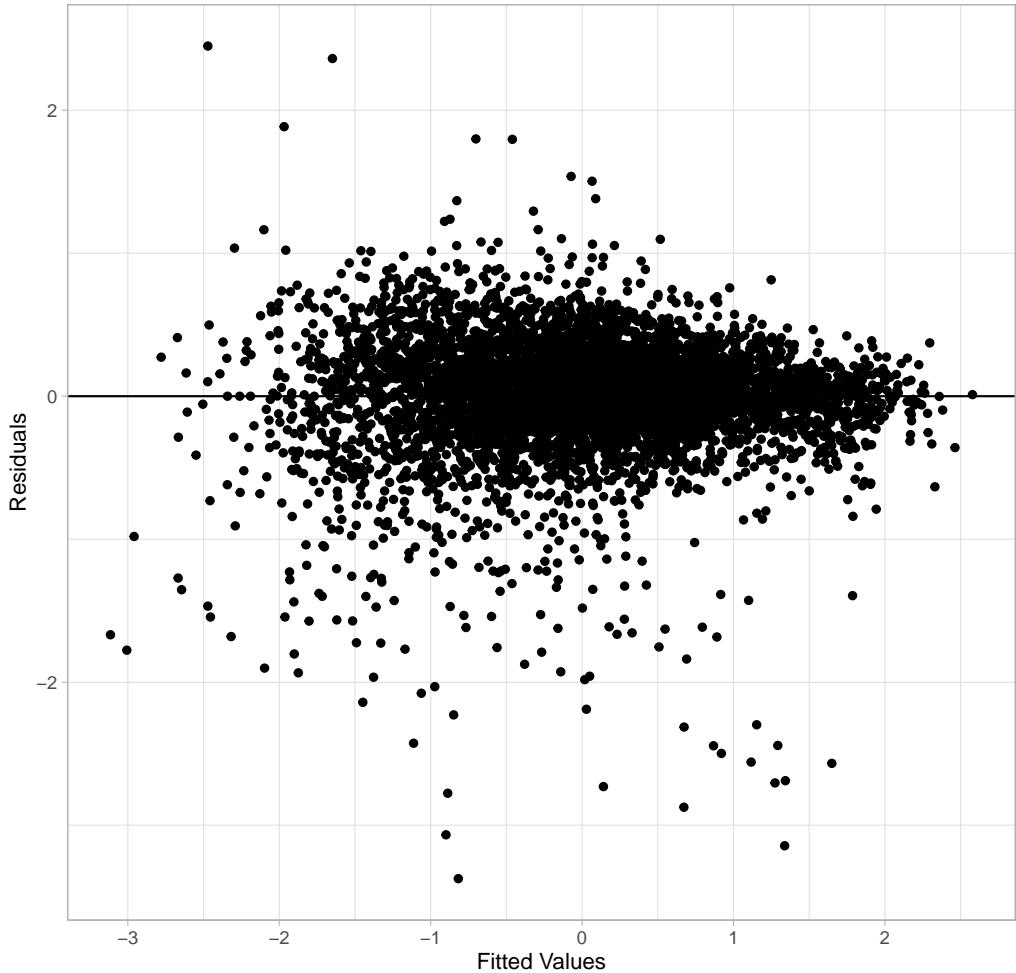


Figure A.8: Residuals vs fitted values for Model 2.

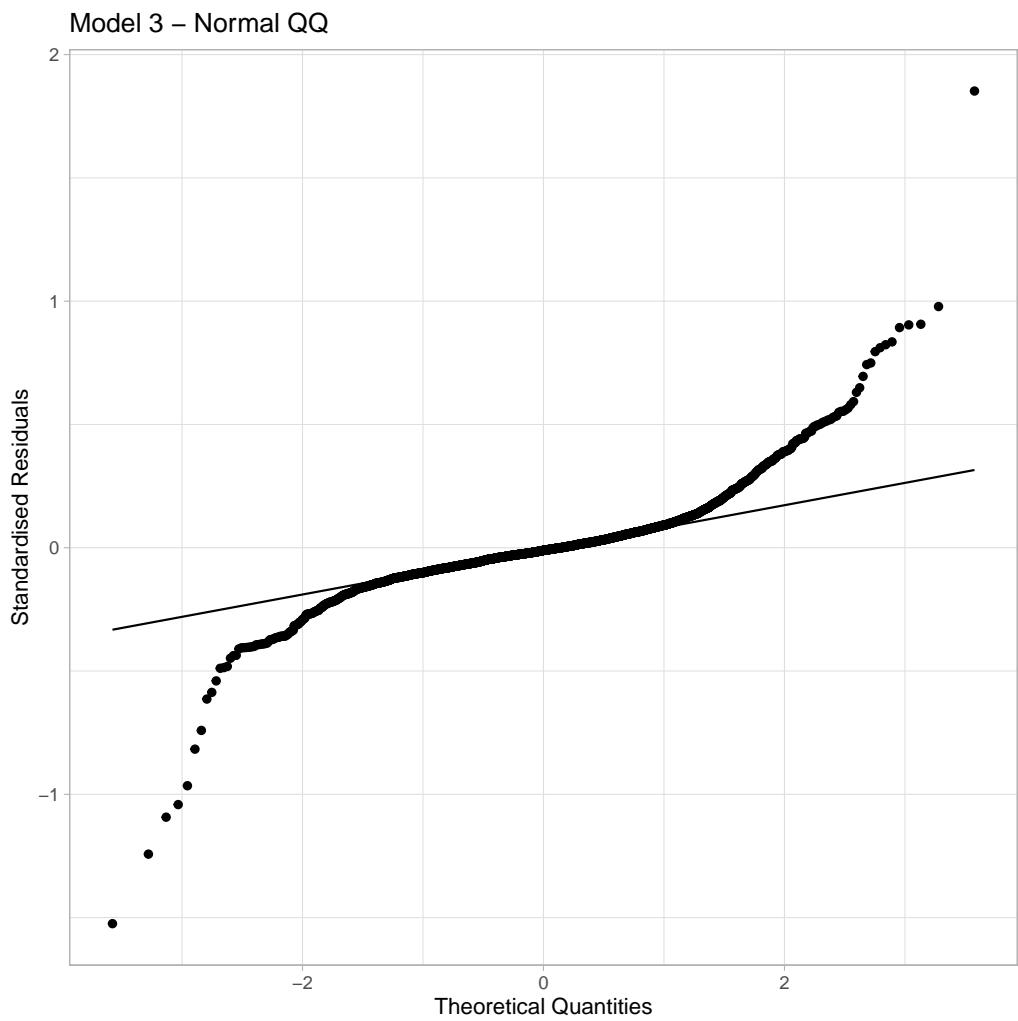


Figure A.9: QQ-plot of Model 3.

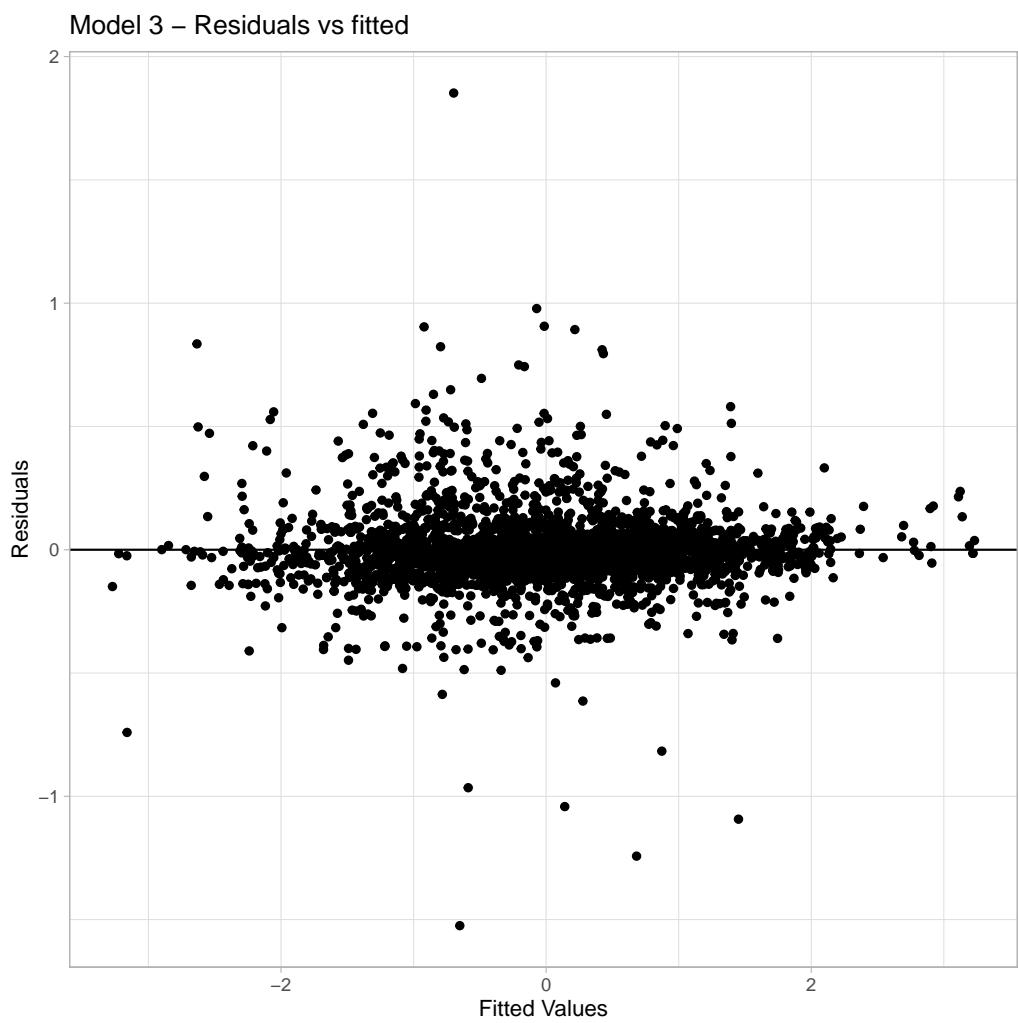


Figure A.10: Residuals vs fitted values for Model 3.

Model 4 – Normal QQ

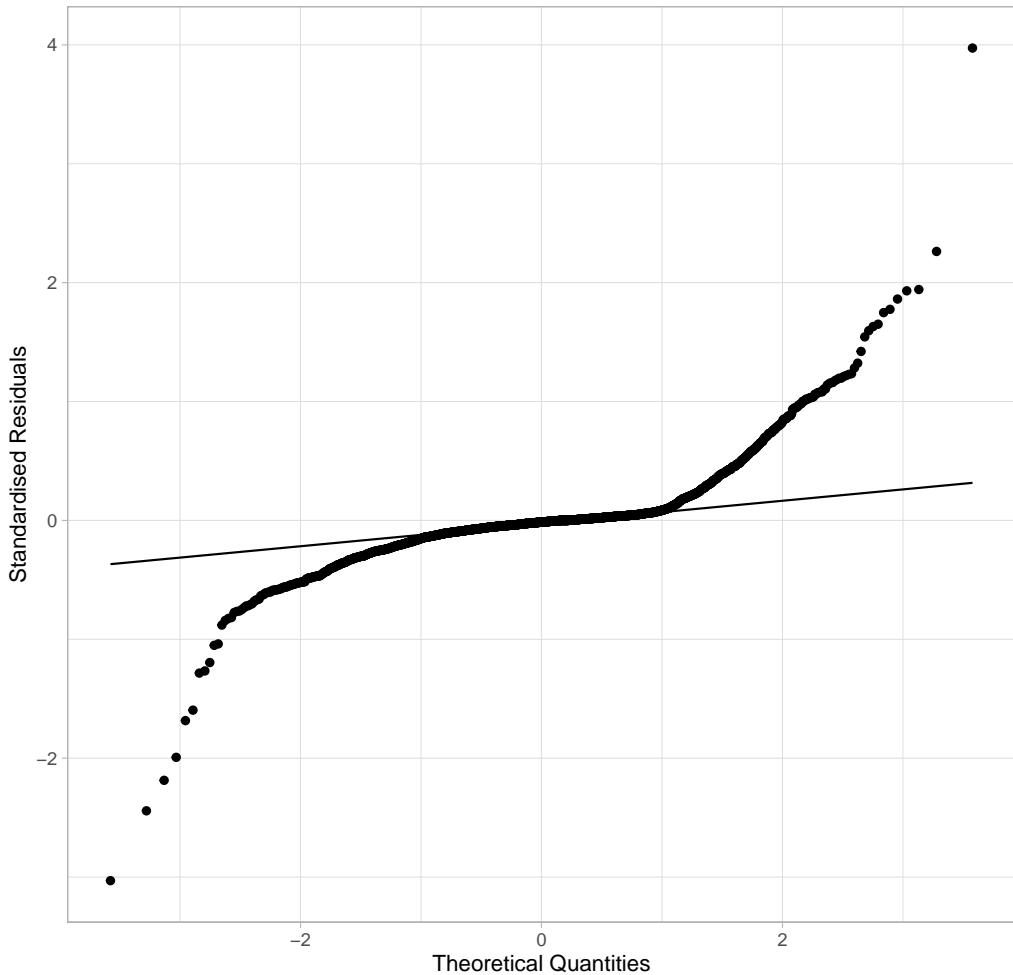


Figure A.11: QQ-plot of Model 4.

Model 4 – Residuals vs fitted

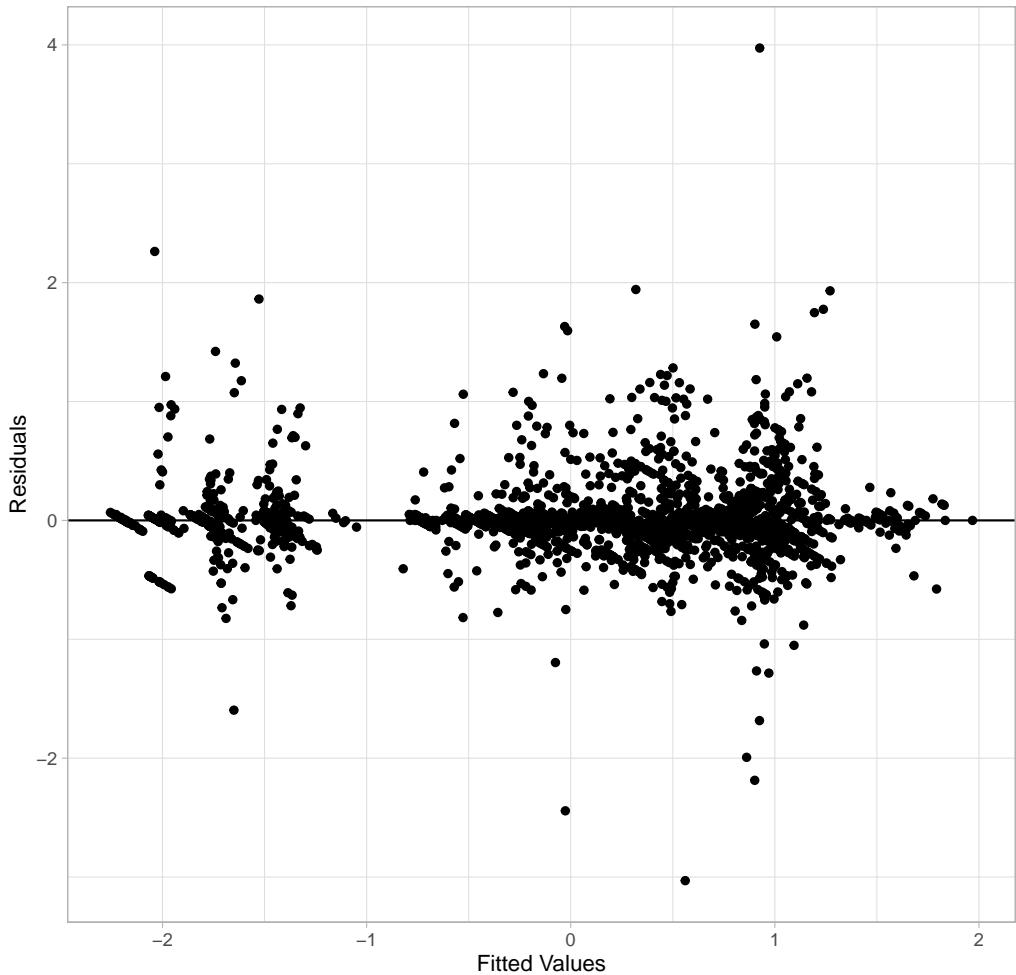


Figure A.12: Residuals vs fitted values for Model 4.

Model 5 – Normal QQ

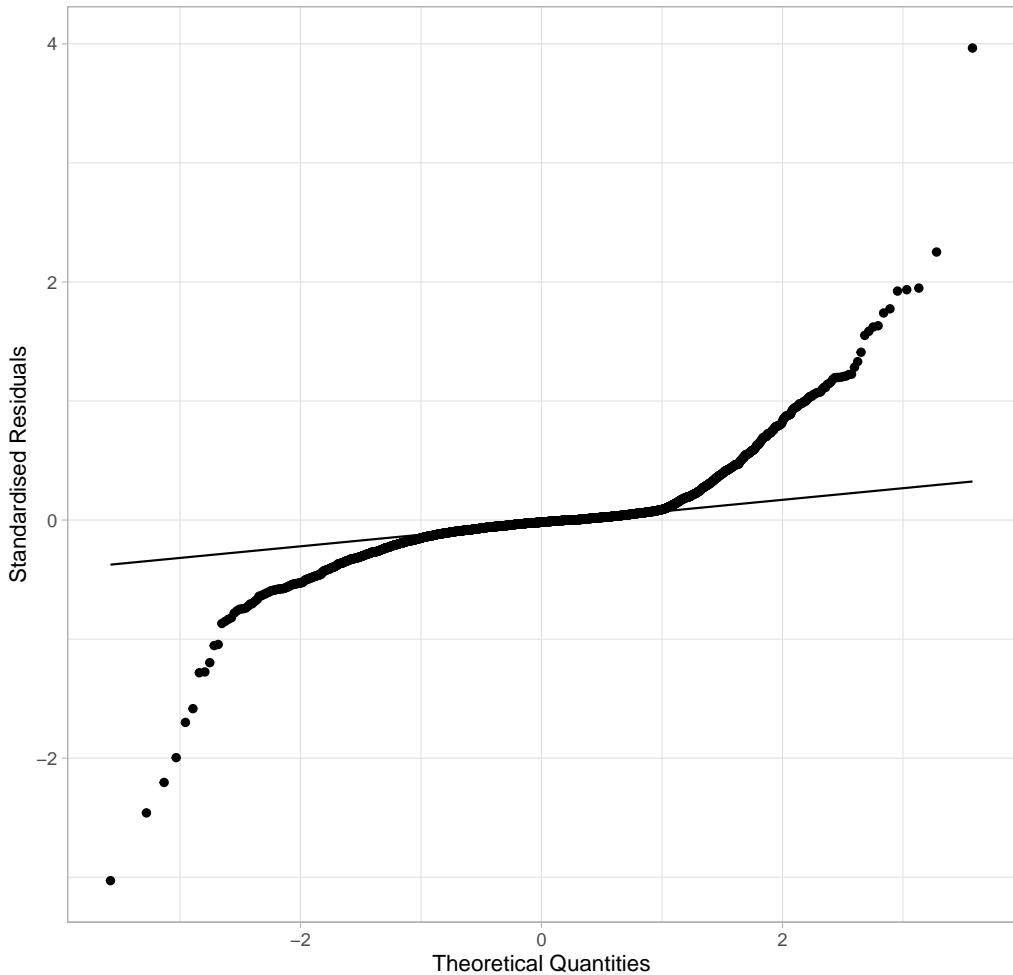


Figure A.13: QQ-plot of Model 5.

Model 5 – Residuals vs fitted

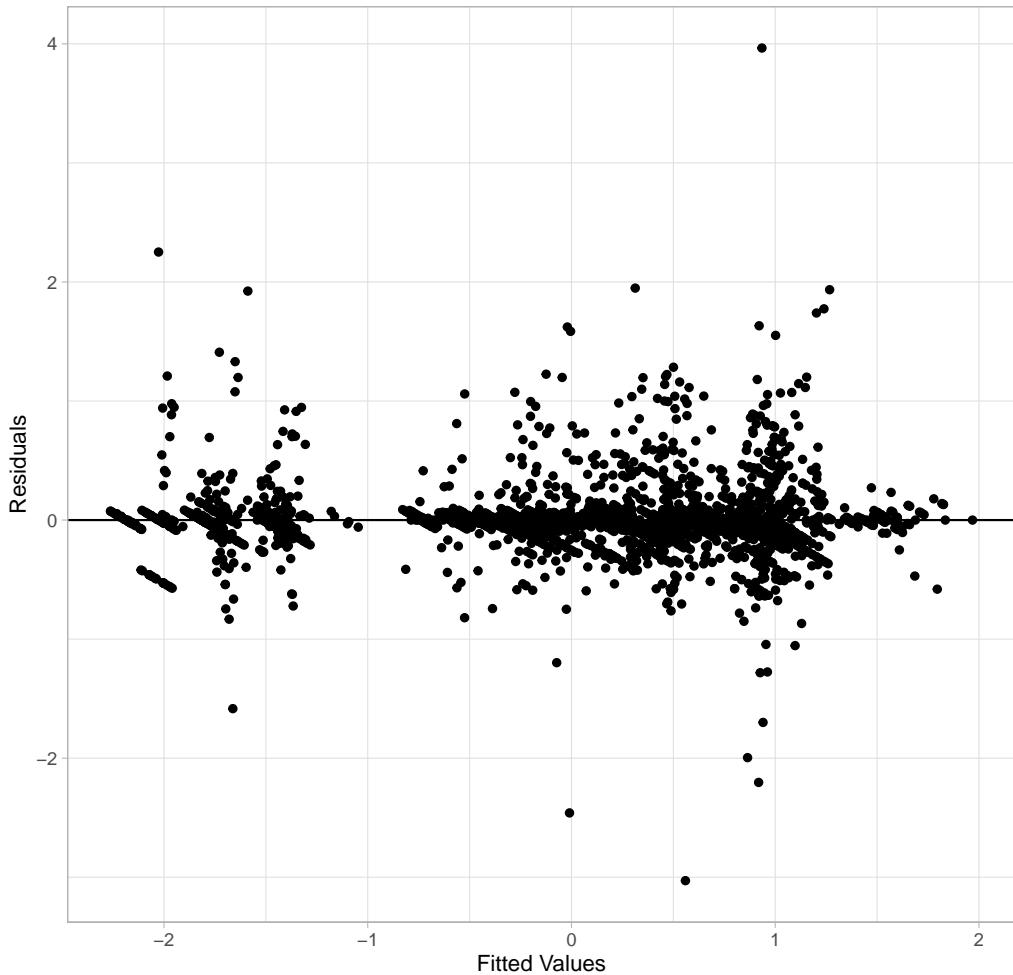


Figure A.14: Residuals vs fitted values for Model 5.

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