iVES OENO One vine & wine

DOI: https://doi.org/10.20870/oeno-one.2026.60.2.8381

OPEN ACCESS

In Vino Analytica Scientia DAVIS 2024 Analytical Chemistry for Wine, Brandy and Spirits

SPECIAL ISSUE

On terroir – The choice of model emphasises different measured attributes in data sets

Leigh M. Schmidtke[1,*], Sijing Li[1], Sue Bastian[2], Keren Bindon[3], Marcos Bonada[4], Paul K. Boss[5], Robert G.V. Bramley[5], Paul R. Petrie[2,4], Cassandra Collins[2]

[1] Gulbali Institute, Charles Sturt University, Wagga Wagga, New South Wales, Australia
[2] School of Agriculture, Food and Wine, Waite Research Institute, Adelaide University, Adelaide, South Australia, Australia
[3] The Australian Wine Research Institute, Glen Osmond, South Australia, 5064, Australia
[4] Crop Sciences, South Australian Research and Development Institute, Waite Campus, Adelaide, South Australia, Australia
[5] CSIRO, Waite Campus, Glen Osmond, South Australia, 5064, Australia

Article number: 8381

*correspondence:
lschmidtke@csu.edu.au

Associate editor:
Dominik Durner

Received:
13 November 2024

Accepted:
13 March 2026

Published:
13 May 2026

This article is published under the **Creative Commons licence** (CC BY 4.0).

Use of all or part of the content of this article must mention the authors, the year of publication, the title, the name of the journal, the volume, the pages and the DOI in compliance with the information given above.

KEYWORDS: IVAS 2024, chemometry, machine learning, viticulture, oenology

This article is published in cooperation with the 13th edition of In Vino Analytica Scientia 2024 (IVAS 2024), 9-12 July 2024, Davis, California.

Guest editors: Andrew Clark, Leigh Schmidtke, and Aurélie Roland.

ABSTRACT

The terroir concept relates to the specific growing conditions of landscape, geography, climate, and, importantly, the interactions of people through farming practices that impact the composition of a primary product and its gustatory characteristics. Defining compositional measures of terroir is challenging because many attributes combine to create a wine's terroir. In the present study, a temporal and spatial investigation comprising targeted measures of viticultural practice, grape and wine composition, and sensory ratings was undertaken across subregions in the Barossa Geographical Indication. Data were arranged into multiple blocks associated with viticultural variables, grape or wine composition, and sensory ratings. Statistical approaches to analysing the data included *k*-means clustering, ANOVA Multiblock Orthogonal PLS (AMOPLS), random forest ensemble (RF), boosted classification decision trees (BCT) and artificial neural networks (NN). Two or three clusters of samples were evident in vintage *k*-means models, and the clusters were correlated with vineyard elevation and temperature accumulation. AMOPLS consistently extracted predictive latent variables associated with five levels for the subregion explanatory factor for vintage models, but predictive scores for subregion were not evident when the entire data set was decomposed in a single model, suggesting that subregions were inconsistent for the expression of vine, grape, and wine composition over the duration of the study. Optimised RF or BCT models performed on par with similar overall classification errors of between 11 and 14 % using a reduced feature set derived through recursive elimination of less important features. Feature importance in the two decision tree ensembles varied slightly, with RF models selecting features from most data blocks and BCT employing more features related to wine composition. The NN model was the least accurate model for sample classification assignment. *k*-means clusters were heavily dependent on measures of grape and wine volatiles with contributions from grape amino acid data. AMOPLS features of importance were distributed throughout the data blocks with viticultural characteristics, grape carotenoids, phenolics, and wine polysaccharides making the largest contributions to model outcomes. These results provide insight into different data modelling approaches, which may provide similar outcomes for classification or clustering of samples, but the influential features within specific models may be different; this will impact the interpretation of the importance of these measures of composition in terroir models.

INTRODUCTION

The association of high-value agricultural products to a specific location that infers distinctive compositional and sensorial qualities that cannot be replicated elsewhere underlies the terroir concept (Brillante *et al.*, 2020; Ceccarelli *et al.*, 2010; van Leeuwen *et al.*, 2004). Terroir is a well-known concept which relates the characteristics of a wine to the specific region in which its grapes were grown and the wine made, and enables consumers to be assured of the wine's style and provenance. Complex interactions between regional climate, topography, plants, animals, soils, and people create a highly woven fabric of connectedness that imparts influence upon products which may or may not be evident to the consumer. Whilst terroir is mostly associated with wine, the concept is equally applicable to a diverse cohort of products for which a unique quality and pedigree is claimed. Appellation control designations may be used to indicate regionality and quality levels of a product, that is, an associated terroir, and this concept has been applied to agricultural products for centuries (Ceccarelli *et al.*, 2010). For example, the French *Appellation d'Origine Contrôlée* (AOC) certifies a product has been produced within a specific geographical region, that it possesses the typicality – *i.e.*, composition, flavour and mouthfeel – expected for produce from the region in which it is produced, and so contributes to its authenticity, thereby providing a guarantee to consumers as well as delivering unique points of difference in the marketplace (Ceccarelli *et al.*, 2010).

The geographical scale for which a terroir may be considered appropriate varies considerably, evident by the size of various geographical indications (GI), the delineation of sub-regions within these, and the fact that variations in terroir may also be expressed at the within-vineyard scale (Bramley *et al.*, 2011; Bramley *et al.*, 2017). Large GIs are likely to have multiple terroirs, and regions may thus be divided into subregions for the purposes of defining typicality. Typicality, or *typicité*, a term generally referring to sensory domains or features that are consistently present in wines from a region may be evident for consumers with familiarity with wines from the region (Ballester, 2021; Schüttler, 2013). As terroir comprises a significant element of human interaction during the grape and wine production processes, it is reasonable to expect that wines from a region will possess elements of consistent sensory expression but may also evolve over time to account for moderations of human input to effect emerging regional styles, and in response to changing climates. Thus, terroir is not a static concept but may evolve as factors important in grape and wine production are altered in response to biotic and abiotic factors or human interventions (Drappier *et al.*, 2019). Using expert panellists familiar with regional wine characters to select exemplars, a clear and distinctive subset of inter-regional sensory profiles of Australian Shiraz was determined, indicating an evolution of terroir across multiple wine regions (Pearson *et al.*, 2020), with some of the identified wine sensory features specifically associated with regional wine styles attributed to measures of climate (Pearson *et al.*, 2021). Given that important regional attributes contributing to wine typicality, and thus terroir, are the cultivar, climate, and soil (van Leeuwen *et al.*, 2004), it follows that measures of climate and soil attributes may provide an indication of subregional demarcation within GIs (Bramley & Ouzman, 2022).

A significant challenge in understanding the impact of specific abiotic and biotic factors, for example, temperature at varying scales, solar radiation, soil properties, water, and nutrient availability, that are inexplicitly associated with vine performance, and which in turn impact wine composition and sensory features, is decoupling confounding experimental factors within the context of experimental observations. Recent reviews that synthesise research findings from multiple investigations present an interpretation of the influence of climate components, vineyard characteristics (van Leeuwen *et al.*, 2020), and agronomic practises (Alem *et al.*, 2019) that influence wine style. Well-designed field experiments have provided insights into the influence of plant genetic and abiotic environmental factors upon the plasticity of berry gene expression (Santo *et al.*, 2018). Investigations to determine the impact of abiotic factors on wine sensory domains have also helped to understand the relative contributions of cultivar and grape harvest dates to wine styles (Antalick *et al.*, 2021; Schmidtke *et al.*, 2020).

Options for modelling data sets of grape and wine composition for understanding the factors that influence terroir are numerous and these may be broadly grouped as linear and non-linear methods. Linear models seek to explain observations through proportional combinations of the measured attributes used for modelling; they are computationally efficient and relatively easy to interpret. Non-linear models are more complex as they seek to explain observations that do not change proportionally to the data measurements used to construct the model, making interpretation and relating changes in measured attributes to the modelled observations more difficult. Biological systems are typically non-linear (Manicka *et al.*, 2023). Unsupervised clustering methods allow the data to reveal patterns of samples without prior knowledge of identity, whereas supervised classification methods seek to discriminate samples based on *a priori* information. Each approach has merit, providing appropriate quality assurance when modeling data. Clustering algorithms may be either agglomerative, in which all samples commence as their own cluster and are iteratively assigned to closely positioned samples, reducing the number of derived clusters, or partitional, in which all samples commence with assignment to a single group and are iteratively segregated into clusters. Either approach requires the computation of a distance measure between samples, which are typically linear combinations of the data, with the goal of sample assignment to a cluster solution that minimises intra-cluster distances and maximises inter-cluster separation. Many examples of clustering in viticulture are presented in the literature, with *k*-means, a linear method, being a popular approach. Supervised data analysis makes use of information about the sample class to extract meaningful information. The AMOPLS framework, which is also a linear method,

has its origins in metabolomics (Boccard & Rudaz, 2016) and has been successfully used to determine the underlying contribution of grape production site, harvest timing, and cultivar with complex data sets (Schmidtke *et al.*, 2020). An advantage of the AMOPLS method is the sequential extraction of information from multiple data sets using explanatory factors (EF) associated with the experimental design, after the extraction of orthogonal information to the factors of interest. This is analogous to data decomposition approaches for ANOVA models and enables robust hypothesis testing and interpretation of the experimental design, the significance of the factor level contributions, and their interactions. Extraction and subtraction of the orthogonal information thereby simplifies model interpretation (Boccard & Rudaz, 2016).

Decision trees are extremely popular non-linear machine learning methods for classification with a shared approach of combining a series of weak learners (decision splits) to create a strong predictive algorithm. The random forest approach (Breiman, 2001) is a bootstrapped aggregation (bagged) ensemble of decision trees in which many noisy and largely unbiased models of weak learners are averaged to create a robust model. For every tree in the ensemble, a random selection of variables can be used in the decision split, and each tree is grown with a fraction of samples left out (out-of-bag), and these are used to determine model error. A process of simultaneous cross-validation and model optimisation is achieved through minimisation of the out-of-bag error (Friedman, 2017). The outcome for a random forest ensemble can be either an average (for regression) of the series of parallel decision trees of randomly selected variables, or, for classification, either taking the majority of votes that the sample of interest received or by averaging the probabilities for class assignment for the samples (Stavropoulos *et al.*, 2020). Random forest models are less prone to overfitting and are robust to noisy data (Stavropoulos *et al.*, 2020), which is common for natural science datasets (Bunde, 2023), making them ideally suited for terroir-related studies (Oliveira *et al.*, 2025).

Boosted ensemble trees fit a series of weak learner decision splits, or stumps, to the training data in a sequential manner to lower the error of prediction, so that each new decision split emphasises the residuals. To begin, all samples are equally weighted, and the decision split is chosen to minimise the predictive error, or loss function. Incorrectly classified samples are reweighted higher than correctly classified samples, and another decision split is added to the model. The emphasis for boosted ensembles is on correcting errors by iteratively and sequentially reweighting samples and building additional decision splits using the misclassified samples, thereby increasingly focusing on the most difficult samples to predict. The final model may consist of linear combinations of many (100–1000s) decision trees (Elith *et al.*, 2008). Several boosting algorithms are commonly used, with multiclass adaptive boosting, AdaBoost (Zhu *et al.*, 2009), being one of the most popular (Stavropoulos *et al.*, 2020). Considerable attention to boosted ensembles has arisen with variations to algorithms to accommodate massive data sets, resulting in varying approaches for minimising loss functions and incorporating deeper decision trees, including up to eight levels in extreme gradient boosting, XGBoost (Chen & Guestrin, 2016), feature selection and bundling approaches in light gradient boosting (Ke *et al.*, 2017), and categorical data. Whilst each approach has advantages, all boosted ensemble methods must be carefully tuned for learning rate, complexity, and appropriately validated using independent sample sets (Elith *et al.*, 2008). Boosted models have recently been applied to wine chemical and sensory datasets with superior modelling outcomes and, therefore, present a promising approach to wine datasets (Sáenz-Navajas *et al.*, 2025).

Artificial neural networks (NN) derive their name through a perceived similarity with biological neurons in the brain. NN models consist of an input layer, being the data of interest, connected to hidden layers of nodes via weighted connections and combinations of the dataset. A network may be one or several hidden layers deep, and the number of nodes, or width, of each layer is chosen during the optimisation process (Aggarwal, 2023). The final output layer has an identical number of nodes determined by the type of application (regression or classification). Once a layer is defined, each connection is assigned a weight, which defines the importance of the connection, and a bias, which may reposition the activation function of each node. Each node within a layer may feed forward through an activation function, which is typically a non-linear transform of the weights and input of that node and the bias. Selective activation of nodes is important since if every node were to feed forward, the layer would resemble a linear combination of the inputs (Marini, 2020). Various activation functions may be used, depending upon the data structure and application, with the rectified linear activation function, which only feeds forward when the activation output is positive, being commonly used. Adjustments to connection weights and biases are made through an iterative process to minimise the error (loss function) and maximise the model fit to the data during the training process. Neural networks are emerging as important modelling algorithms for food authenticity and traceability studies (Li *et al.*, 2025; Ma *et al.*, 2024) and may outperform linear discrimination models for wine classifications (Cosme *et al.*, 2021).

A major project conducted in the "Barossa Zone" GI of South Australia sought to understand the compositional and sensorial basis of terroir variation in Shiraz wines (Schmidtke *et al.*, 2024) and its alignment to regional scale variation in landscape (soil, topography) and climatic factors (Bramley & Ouzman, 2022). Of particular interest was the alignment of variation in these factors to a delineation of subregions within the GI proposed by the local wine industry. To address the latent multifactorial nature of terroir, multiple machine learning algorithms – each with different strengths, assumptions, and interpretative frameworks – were applied. Irrespective of the data modelling approach, different algorithms may place differing emphasis on the measured

attributes, even if a consistent measure of feature importance is used (Wang *et al.*, 2024). Rather than undermining the value of the models, this variation reflects the complexity of terroir and underscores the need to contextualise model outputs within the current understanding of viticultural and environmental interactions. This approach does not aim to reaffirm heuristic knowledge but to use data-driven methods to challenge, refine, or validate it, providing a framework for future empirical exploration of terroir-related phenomena. Using the same data set from the Barossa, the purpose of this present investigation is to explore how common linear and non-linear modelling approaches vary in output and emphasis for measured attributes, thereby influencing the interpretation of viticultural factors upon the terroir concept.

MATERIALS AND METHODS

1. Experimental design and datasets

The Barossa Zone GI comprises approximately 14,000 ha of vines within two regions, namely the Barossa Valley and Eden Valley, and a recognised subregion within the latter, High Eden (Wine Australia, 2023). Wine sensory assessments undertaken by Barossa vignerons (www.barossawine.com/vineyards/barossa-grounds/) over numerous vintages have been used to informally designate additional subregions to the Barossa Zone, and these subregions are designated Northern Grounds (NG), Central Grounds (CG), Eastern Edge (EE), Southern Grounds (SG), Western Ridge (WR), and Eden Valley (EV). Four sites from each subregion were established in the 2017/2018 growing season, with an additional site selected in EE during 2018/2019 to substitute one site that became unavailable after the first year of the study (Figure 1). Three sampling zones for each site were identified based upon standard geospatial measures (Bramley *et al.*, 2011) to capture site heterogeneity. Analytical workflows for vine performance, crop maturation, grape composition, and wine composition and datasets were collated into distinct data blocks (Table 1) as comprehensively described by Schmidtke *et al*. 2024.

A consensus sensory space for each vintage was determined using STATIS (Abdi *et al.*, 2012) to accommodate differences in the sample significance of sensory adjectives across the study period and high-dimensional sensory datasets. From each of the consensus multivariate spaces created for each vintage sensory data set, five latent variables representing 18–20 % of total data variance (Figure S1) for each sensory data set from each vintage were used as a surrogate sensory data block. These five latent variables comprised the 15th data block for multiblock clustering and discriminant analysis.

Data were arranged as 15 discrete data blocks containing varying numbers of attributes associated with specific measures of viticultural performance, grape and wine composition, or latent variables (LV) summarised in Table 1. Where possible, each model was constructed using the largest possible number of site-sample replicates and measured attributes, and a more in-depth description of each data block and attribute selection for modelling is presented in Table S6.

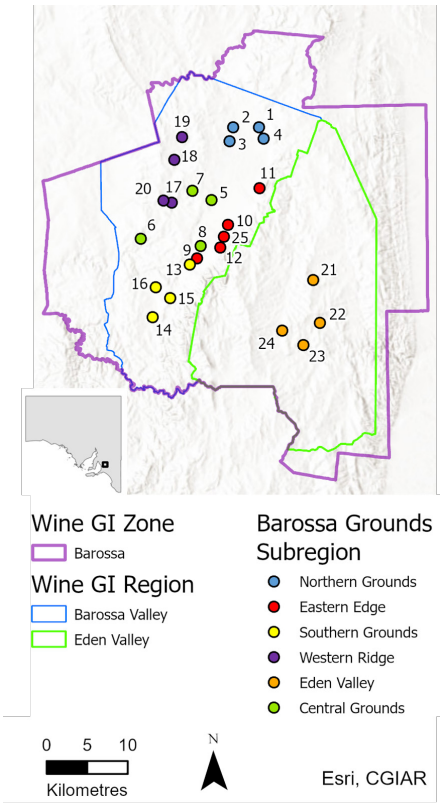


FIGURE 1. The Barossa Zone geographical indication comprises the defined regions of Barossa Valley and Eden Valley. The approximate location of vineyard trial sites and their nominal subregion.

TABLE 1. Data block description including colour code used in data decomposition, block order and number of measured attributes for each table are indicated in parentheses.

Block #	Description
1	Vine performance (7)
2	Crop maturation (12)
3	Grape amino acids (22)
4	Grape bound volatiles (106)
5	Grape carotenoids (14)
6	Grape fatty acids (29)
7	Grape volatiles (101)
8	Grape phenolics (19)
9	Grape tannin (16)
10	Wine basic chemistry (10)
11	Wine metals (15)
12	Wine phenolics & colour (51)
13	Wine polysaccharides (11)
14	Wine volatiles (63)
15	Wine sensory (5 latent variables)

2. Sample selection for data modelling

Samples for individual vintage and combined vintage chemometric models were selected based on the availability of full data sets for each sample site replicate across each of the 15 data blocks. Site and replicate sample selection for each vintage is presented in Table S7, noting that combined vintage models comprised a concatenation of each of the vintage sample data sets. Where data for a measured attribute was missing for a specific sample replicate, median substitution for a value was used based upon the sample set, with a maximum of 0.05 percent of data substitution to complete the vintage data set. No outliers were evident upon initial inspection of the data; therefore, a standardised approach to data scaling, unit scaling, was used in all models to maintain consistency across data decomposition methods.

2.1. Discriminant analysis – AMOPLS models

ANOVA Multiblock Orthogonal Partial Least Squares analysis (Boccard & Rudaz, 2016) of the datasets was undertaken using methods for viticulture and wine analysis (Schmidtke *et al.*, 2020). AMOPLS for all vintages with Explanatory Factors (EF) *Vintage* and S*ubregion*, and single vintages with EF *Subregion* and *Site* were modelled without interaction. The ANOVA framework requires a consistent selection of variables across the sample sets, and, therefore, measured attributes were selected that span the data and sample range for each vintage (2018 to 2020), which means that some variables (Table S6) and some *Subregion* site replicates (Table S7) were not modelled in each year. Removal of orthogonal variance in these models was tested using 1000 × permutations of the samples within each explanatory factor to determine the appropriate numbers of the extracted components for each model. The contributions of the explanatory factor to the remaining dataset once orthogonal variance was subtracted were then modelled, and the overall goodness of fit was determined from the residual structure ratio (RSR). Overall data block contributions to AMOPLS predictive scores associated with explanatory factors and levels were determined from the corresponding predictive scores and association matrices, and the contribution of specific variables to each model explanatory factor was determined by variable importance to projection (VIP) for multiblock models (González-Ruiz *et al.*, 2017).

2.2. Unsupervised clustering – *k*-means

Unsupervised clustering of samples was undertaken using *k*-means clustering using the same measured attributes used for AMOPLS discriminant analysis. Prior to *k*-means clustering, each data block for each vintage model was mean-centred and variance scaled to unity before concatenation to create a data superblock, which was passed to the *k*-means algorithm. Each variable in all data blocks, therefore, contributed equal variance. To overcome vintage influences for the all-vintage *k*-means cluster, a combined vintage data superblock was constructed by concatenation of the mean-centred and variance-scaled attributes in each data block for each vintage. Each vintage of data, therefore, contributed the same variance to the data superblock, and each measured variable had equal variance for each year. *k*-means clustering was undertaken using the data superblocks with distance measures between samples using the "cityblock" approach, as this metric provides a higher contrast than Euclidean distances for data dimensionalities larger than 20 (Aggarwal *et al.*, 2001). *k*-means clustering was repeated with five replicates for each iteration and between two and six clusters were chosen for testing. The appropriate number of clusters for each model was based upon the inspection of silhouette profiles (Rousseeuw, 1987) to identify a stable clustering outcome of the models. As the *k*-means approach commences with a random sample selection, the outcomes from *k*-means models were inspected and two pairs of samples that were consistently modelled together were used as anchors to ensure a consistent presentation of the site in multiple models. Samples from NG (sites 1 and 4) and EV (sites 23 and 24) were anchored to cluster 1 and 2, respectively, and cluster assignments for other sites recoded for consistent presentation of cluster labels for each model. A measure of variable contributions to the *k*-means solutions was determined by assessing the ratio of variable sums of squares for in-cluster samples to the entire variable table sums of squares. Correlation of clusters to vineyard site parameters (elevation, cation exchange capacity, average water holding capacity, 30-year mean (to 2018) growing degree days, mean January temperature, mean growing season temperature, growing season rainfall and annual rainfall with the *k*-means solutions was done based upon the range of observed values and assigning each site to one of three bands (low, medium or high) with boundaries based upon the 33rd and 66th percentiles of observations.

2.3. Bagged ensemble classification tree – random forest

Ensembles of random forest (RF) decision trees (Treebagger, MATLAB) were constructed from the dataset used for the all-vintage *k*-means clustering with a random stratified partitioning of the data to training and independent test sets at a ratio of 70/30 of samples, respectively, which approximates the default settings (1/*e*), producing unbiased error estimates (Breiman, 2001). An iterative process comprising optimisation of the number of grown trees [25 50 75 100 125] and minimum leaf size [1 5 10 20 50 100 150], followed by removal of features with low contribution to prediction outcomes, was undertaken. For each iterative model, the number of predictors to sample for each decision split was set to the square root of the total number of features. Selective elimination of features was undertaken by recursively removing ten percent of features with low predictive importance determined using the mean absolute Shapley values for all query points in the out-of-bag sample set. The model with the lowest overall out-of-bag error prediction was selected for measuring predictive acuity by computing area under receiver operator curves (AUROC) for the training and independent test data set with 95 % confidence intervals determined from 100 bootstrapped iterations, and a confusion matrix for class prediction (1 against all) for all samples in the independent test set. A proximity matrix defined as the fraction of trees in the ensemble for which two observations land on the same leaf for samples was

determined and used for multidimensional scaling (MDS), and class boundary confidence intervals were determined with Hotelling's T^2. Feature importance in the final model was determined using Shapley values for all query points using the independent test set.

2.4. Boosted ensemble classification trees

Ensembles of boosted classification trees (BCT) (AdaBoost. M2, MATLAB) were trained with a stratified partition of samples drawn from the same dataset for *k*-means clustering at a ratio of 70/30 for training and test sets. Each ensemble was trained for learning rate [0.001 0.1 0.25 0.5 1], maximum data splits [1 3 9 15] and the number of trees [10 50 120 150 200] using a grid search and cross-validation with 10 data splits to determine model error at each grid location. The model with the minimum cross-validated error was used to determine the importance of features using mean absolute Shapley values at all query points in the training set. Removal of ten percent of features with low predictive importance was done, and a recursive approach to ensemble optimisation with the remaining features was undertaken for 30 iterations. The most parsimonious optimised model with the minimum cross-validated error was used to predict sample class, AUROC with bootstrapped confidence intervals and a confusion matrix (1 against all) for the independent training set. Feature importance for comparison with other models was determined using Shapley values for all query points using the independent test set.

2.5. Artificial neural networks

An optimised neural network was constructed (fitcnet, MATLAB) using a stratified partitioning of training samples comprising 70 % of data from the same dataset for *k*-means clustering. NN models were trained to optimise the number of hidden layers [1 to 3]; hidden units per layer [1–300] and activation function [ReLU, Tanh and sigmoid] using cross-validation with 10 data splits to determine model error for each parameter. The model with the minimum cross-validated error was used to test the overall predictive acuity using the independent test set of samples with bootstrapped confidence intervals for AUROC and a confusion matrix for independent test set prediction outcomes. Variable importance for the selected model was determined using mean absolute Shapley values for every query point in the independent test set, and the most influential variables were plotted for the sample class. All data modelling was conducted using MATLAB Version: 24.1.0.2603908 (R2024a) Update 3 (The Mathworks, Natick, MA).

RESULTS AND DISCUSSION

1. Comparison of model performance and insight into terroir

The contribution of each data block to each model tested is illustrated in Figure 2. Differences in the selection and importance of features from data blocks for each model are readily apparent, with marked variations in the features selected for each model relative to the data block contribution to the entire data set. In the present investigation, data were normalised for year and data block to ensure a similar weighting of variables in each model, and thus the size of each data block does not influence the model outcomes.

Notably, the *k*-means cluster solutions have influential measures of grape volatiles for sample cluster assignment at a rate exceeding the overall representation of this data block in the data set. Over-representation of data blocks, although at a smaller magnitude, is also evident for wine volatiles and grape amino acids for the *k*-means cluster sample assignments. The number of influential features in data blocks for vine performance, crop maturation, grape carotenoids, and wine metals is similar to the overall data block contribution to the data set, with other data blocks under-represented.

In contrast to the *k*-means clusters, the AMOPLS models have no significantly contributing features from the grape volatiles, grape-bound volatiles or wine phenolic and colour data blocks for modelling *Subregion*. The importance of data blocks for measures of vine performance, crop maturation, grape carotenoids, grape phenolics, wine basic chemistry and wine polysaccharides is evident, with other data blocks represented at levels similar to the block data set (grape tannin, wine metals and wine sensory) or under-represented. Contributions from the vine performance and crop maturation data blocks in the AMOPLS models suggest differing vineyard management approaches across subregions, potentially to accommodate differences in Growing Degree Days (GDD) and water availability, which impact vine fruit production but not specifically crop composition.

Some interesting comparisons and contrasts between the RF and BCT models for feature importance are apparent. The selected RF model with the lowest overall error comprised 88 features (Figure 3A). However, further reduction of features did not consistently decrease model performance until the number of features for modelling data was less than 40, occurring at the 26th iteration of feature removal, suggesting that many of the features in the chosen model contribute only marginally to the *Subregion* sample class assignment. The optimised RF model comprised 100 trees and a minimum leaf size of 1. Influential features from the vine performance, grape fatty acids, grape phenolics and wine metals data blocks are over represented relative to data block size in the data set for the RF model, with the most influential feature being diethyl succinate, which is also the most influential feature in the BCT model. Under represented data blocks for the RF model are grape bound volatiles, wine phenolics and colour and wine volatiles; data blocks representing measures of grape carotenoids and grape volatiles are represented at levels approximating the proportion of data blocks size in the data set. No features of influence are present from the crop maturation, grape amino acids, grape tannin, wine basic chemistry, wine polysaccharides, and wine sensory data blocks for the RF model.

The BCT model with the lowest overall error (Figure 4A) comprised 55 features for classification, and the removal of too many features resulted in a rapid increase in model error. The optimised model consisted of 200 trees, 15 data

splits and a learning rate equal to 1. Over-represented data blocks for BCT influential features are vine performance, grape carotenoids, grape phenolics and wine metals; grape fatty acids, grape volatiles, wine basic chemistry and wine polysaccharides are represented at a level consistent with data block contribution to the data set. Features from crop maturation, grape amino acids, grape tannin and wine sensory data blocks are not influential in the BCT model.

Intriguingly, feature importance in the optimised NN model (Figure 5) is considerably different to the RF and BCT models, even though the same measure of feature importance was used, and this likely reflects the highly tuneable nature of the weighting and biases between input variables and nodes in the hidden layers that occurs during the optimisation process of model testing, effectively becoming a feature selection process. Grape-bound volatiles, wine metals, vine performance, grape tannin and wine sensory data blocks are significantly over-represented as features of importance in the NN model, whereas crop maturation, grape amino acids and grape phenolics data blocks do not have any features of influence. Grape fatty acids, wine phenolics and colour, and wine volatiles are under-represented, and the remaining data blocks are represented at levels similar to the overall data block proportion within the dataset.

2. Discriminant analysis – AMOPLS

A significant advantage of the AMOPLS method is the sequential extraction of information from multiple data sets using explanatory factors (EF) within the experimental design, which occurs after the removal of structured information that is orthogonal to the EF of interest. Extraction of the orthogonal information thereby simplifies model interpretation, and the significance of the factor contributions can be interpreted in a way similar to classic ANOVA reports. Furthermore, the order in which predictive scores in each AMOPLS model are sequentially extracted provides an indication of the magnitude of differences for levels within the EF. The first predictive scores describe the dataset in the most robust manner, with higher-order latent variables having diminished predictive acuity. The ratio of residual variance to the main effects for each predictive component (RSR) is a multivariate measure of the level of significance for a specific EF, with larger values indicating greater significance. The relative sums of squares for each EF indicate the total variance captured within that EF for the

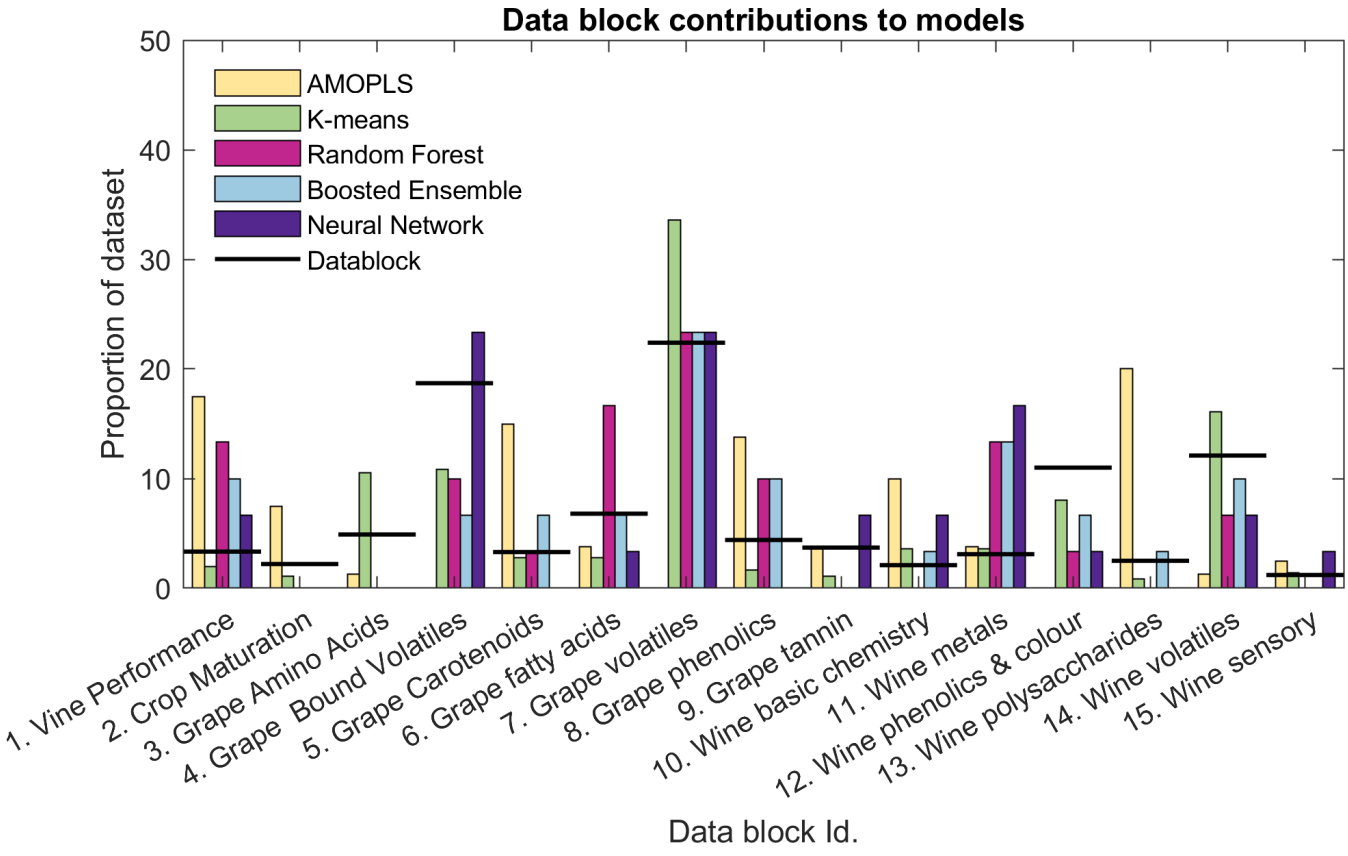


FIGURE 2. Representation of data block contributions and data block size for model performance for subregion discrimination or sample clustering.

Vertical bars represent the proportion of data block contributions based on feature importance measures for each model, with features assigned to blocks according to Table 1 and Table S7. Each horizontal line represents the proportion of the individual data block in the entire data set. Vertical bars above a horizontal line infer that a larger proportion of attributes from the corresponding data block are important in the corresponding model. Vertical bars below a horizontal line imply that a data block is less important relative to the overall size of the data set.

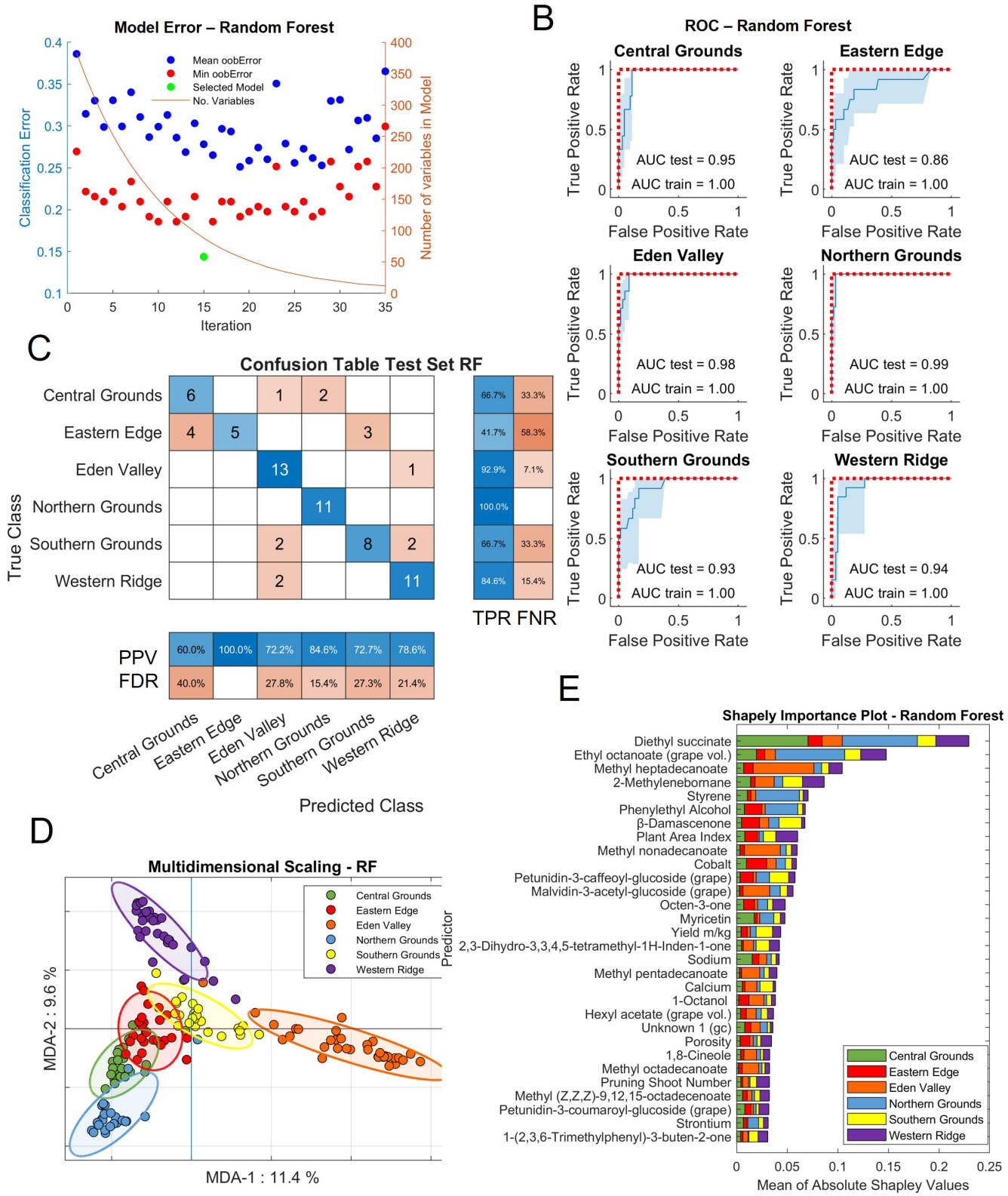


FIGURE 3. Random forest ensemble of decision trees trained for subregion levels using bootstrapped samples.

(A) Recursive reduction in the number of features for each model reduces out-of-bag sample error to a local minimum and then the predictive error increases. The selected model with the lowest overall error (green) was used for assessing receiver operator curves (B) with training (red) and test samples (blue) with bootstrapped confidence intervals for each class and the area under each curve noted. A confusion matrix (C) for test sample prediction shows samples from the Northern Grounds are predicted with the highest true positive rate (TPR); Eastern Edge has the highest positive predictive value (PPV) and the highest false negative rate (FNR); and Central Grounds samples have the highest false discovery rate (FDR). Multidimensional scaling (D) with ellipses representing 95 % confidence intervals. The 30 highest contributing variables to model prediction based on Shapley values are illustrated (E) with bar colours representing the relative importance of features to Subregion class prediction.

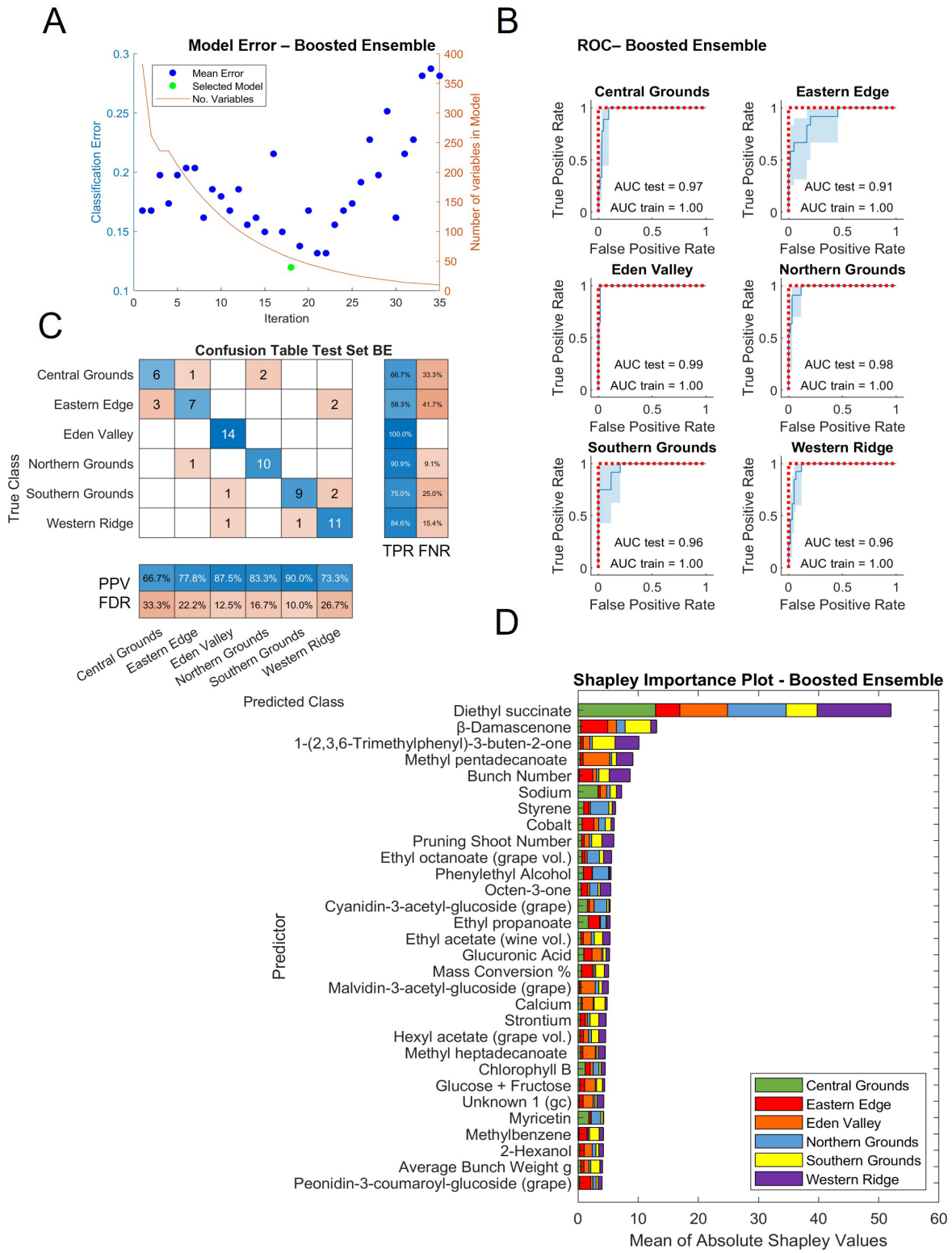


FIGURE 4. Gradient boosted ensemble of decision trees (BCT) trained for subregion levels using bootstrapped samples derived from a 70/30 partition of all samples to training and test sets.

(A) recursive reduction in the number of features used in each model results in a model with a local minima, and then predictive error increases. The model with the lowest overall error (green) was selected for the prediction of subregion class using the independent test set of samples. Model receiver operator curves for training (red) and test (blue) sets with bootstrapped confidence intervals for each class (B), with the area under each curve noted. A confusion matrix (B) for test sample class prediction shows samples from the Eden Valley are predicted with the highest true positive rate (TPR); Southern Grounds samples have the highest positive predictive value (PPV); Central Grounds samples have the highest false discovery rate (FDR), and Eastern Edge samples have the highest false negative rate (FNR). The 30 highest contributing variables to model prediction based on Shapley values are illustrated (D) with bar colours representing the relative importance of features to Subregion class prediction.

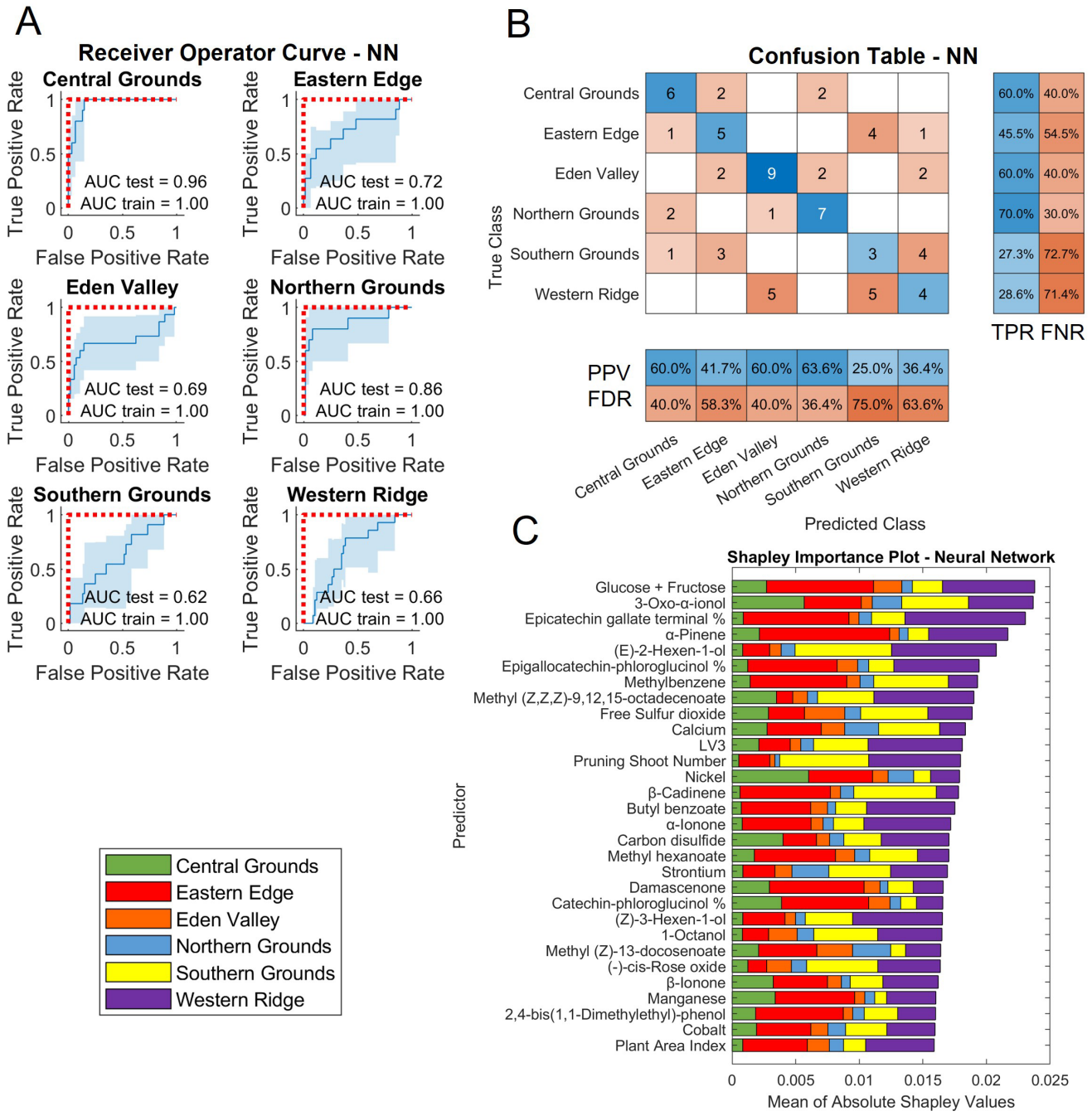


FIGURE 5. An artificial neural network was trained for subregion levels using a 70/30 partition of all samples to training and test sets.

Model receiver operator curves for training (red) and test (blue) sets with bootstrapped confidence intervals for each class, with the area under each curve noted (A). Confusion matrix (B) for test sample class prediction shows samples from the Northern Grounds are predicted with the highest true positive rate (TPR) and highest positive predictive value (PPV); samples from the Southern Grounds have the highest false discovery rate (FDR) and false negative rate (FNR) and (C) the 30 most importance features for the neural network model predictive outcomes based on Shapley values with coloured areas to indicate their relative contribution to subregion classification.

ANOVA model. Overall model performance and significance of explanatory factors based upon permutation testing are presented in Table S8. Significant variance can be attributed to EF *Vintage* (All Vintage model, Residual Structure Ratio (RSR) 2.70, EF *p*-value 0.0001); *Subregion* (All & single vintage models, RSR 1.16-2.20, EF *p*-value 0.0001) and *Site* (Single Vintage Models RSR 2.11–2.35, EF *p*-value 0.0001).

In the "All Vintages" model (EF "*Vintage*" and "*Subregion*"), predictive scores associated with levels within *Vintage* are clear. However, levels for EF *Subregion* are not well discriminated (Figure 6). The first three predictive latent variables (tp1–3) are associated with vintage discernible in the scores plots (left) with data block contributions (right). Sample scores that were colour-coded for GDD (Table S9) show the correlation of the first latent variable (tp1)

with GDD with the discrimination of warm (2018 & 2019) and cool (2020 & 2021) vintage conditions. Of interest is the relative contribution of data blocks 7 and 12 (grape volatiles, wine phenolics & colour) to the dominant "*Vintage"* latent variable. Such observations are well supported in the literature that demonstrates plasticity of expression for grape phenolics and volatile components according to vintages (Antalick *et al*., 2021; Schmidtke *et al*., 2020). GDD was not correlated with any higher-order latent variables associated with either the explanatory factors "*Vintage"* or "*Subregion"* for the four vintage data sets. Common vintage variation for 2019 and 2020 is evident in score plots for tp2; however, the source of this variation is not easily identified and is not related to heat accumulation during the growing season. This latent variable is strongly influenced by grape composition, including maturation, grape volatiles, fatty acids and phenolics (blocks 2, 4, 6, 7, & 8), which also translate into influential wine composition of phenolics and wine volatiles (blocks 12 & 14). Unique vintage variation is evident in tp3 with grape compositional data (blocks 4, 5, 6, 7, 9) and wine phenolics and colour and wine volatiles data largely contributing.

Subregional discrimination in the "All Vintage" model is unclear, although the Eden Valley samples emerge as being somewhat different from other samples used in this model. Higher-order latent variables were not clearly associated with specific subregions (not shown), and a clearer indication of subregion emerges in the single vintage AMOPLS models. Interestingly, there is no significant contribution of the sensory data sets to the extracted scores plots associated with vintage, and sensory data makes a minor contribution to the sample scores for the latent variable associated with subregion (tp4), suggesting sensory differences are consistently present in wines from the Eden Valley compared to other subregions. The relative contribution of each data block to the predictive scores plot for region is heavily skewed to grape composition and wine volatiles and indicates the impact of differences in climate and soil properties between Eden Valley and other subregions (Bramley & Ouzman, 2022) and their consequences for vineyard management, on the differential accumulation of grape components. Subregional discrimination is more clearly evident in the single vintage AMOPLS models (Figures S2–S5), and contributions of each data block to the predictive scores are inconsistent across each year of the investigation.

3. Unsupervised cluster analysis – *k*-means

The results of *k*-means clustering of the replicate observations using the measured attributes for AMOPLS are shown in Figure 7A along with cluster correlations to vineyard attributes. *k*-means cluster solutions were inconsistent with two or three clusters being the optimised solution depending on vintage. In comparing the cluster solutions for *Subregion* site replicates and vineyard characterisation, a general trend consistent with cluster assignment and increasing vineyard elevation is apparent. A correlation to cluster assignment is also evident for soil cation exchange capacity, available water holding capacity, growing degree days, and mean January temperature, with these measures of vineyard decreasing. However, correlation does not infer causation, and the trend observations between vineyard characters and cluster assignment are noted without attribution. A two-cluster solution was appropriate when all vintages were modelled together, and for single vintage models, 2018 and 2020. Eden Valley samples consistently clustered as a group, which also included Southern Grounds (SR14, V18, V20, & V21 [mixed site replicates]) and Eastern Edge (SR25, V20, & V21 [mixed site replicate]) and Northern Grounds (SR2, V20, & V21 [replicate 2]). A three-cluster solution was consistent for vintage 2019 in which one Northern Grounds site (SR3 [replicates 1–3]), Central Grounds (SR5 & 6 [replicates 1–2]; Central Grounds (SR7 & 8 [replicates 1–3]), Western Ridge (SR18 & 20 [replicates 1–3]), Southern Grounds (SR16 [replicates 2–3]), Eastern Edge (SR9 [replicates 1–2]) and Southern Grounds (SR15 [replicate 1]) were grouped together. The three-cluster solution for vintage 2021 is inconsistent with the 2019 three cluster solution with samples from Western Ridge (SR17 [replicate 2]); Western Ridge (SR19 [replicates 1–3]) and Western Ridge (SR20 [replicate 3]) clustering with Southern Grounds (SR14 [replicate 2]); Eastern Edge (SR25 [replicates 1–2]) and Eastern Edge (SR11 [replicates 1–3]). Inconsistent clustering of *Subregion* site replicates implies that variances within the data extraneous to the experimental design contribute equally or greater than inherent biological variation at those sites. Sources of extraneous variation may include analytical performance or remain latent within the dataset. Features with high importance for aggregation of samples to cluster 1 (Figure 7B) include sulfur containing (methionine), branched (leucine, isoleucine, and valine), non-polar (phenylalanine and alanine), polar (serine, glutamine, and asparagine) and positively charged (histidine and arginine) amino acids, grape volatiles (bound and free) and wine volatiles including branched acetate esters (2-methyl propyl acetate, 3-mercaptohexy acetate, 2-methyl butyl acetate, 3-methyl butyl acetate, and 2-ethyl furan). Conversion of specific amino acids to a pool of straight-chain esters, branched esters, and higher alcohols through microbial transformation is well documented (Antalick *et al.*, 2015; Fairbairn *et al.*, 2017; Sumby *et al.*, 2010). Also contributing to cluster 1 sample aggregation is the concentration of 1,8-cineole in grapes and the wine counterpart eucalyptol, and p-xylene in grapes. Xylene is a contaminant in foods and beverages arising from proximity to environmental sources associated with petroleum products and usage (Hinwood *et al.*, 2006), and intriguingly has been isolated with other volatile markers of petroleum pollution (BTEX) from tank headspaces during wine fermentations and transfers (Sanjuán-Herráez *et al.*, 2014). BTEX compounds are lipophilic and have affinity for cuticular wax and plant stomata (Yu *et al.*, 2022), and it is, therefore, reasonable to expect absorption of BTEX volatiles onto grape berry surfaces could occur prior to or during grape harvest. Features of importance for aggregation to cluster 2 (Figure 7B) include C6 volatiles (1-hexanol, (E)-2-hexen-1-ol, 2,4-hexadienal, (E)-2-hexen-1-ol, (Z)-3-hexen-1-ol), volatile acids, terpenes, aldehydes, grape phenolics, and wine metals (zinc, sodium, and magnesium). These results demonstrate the importance

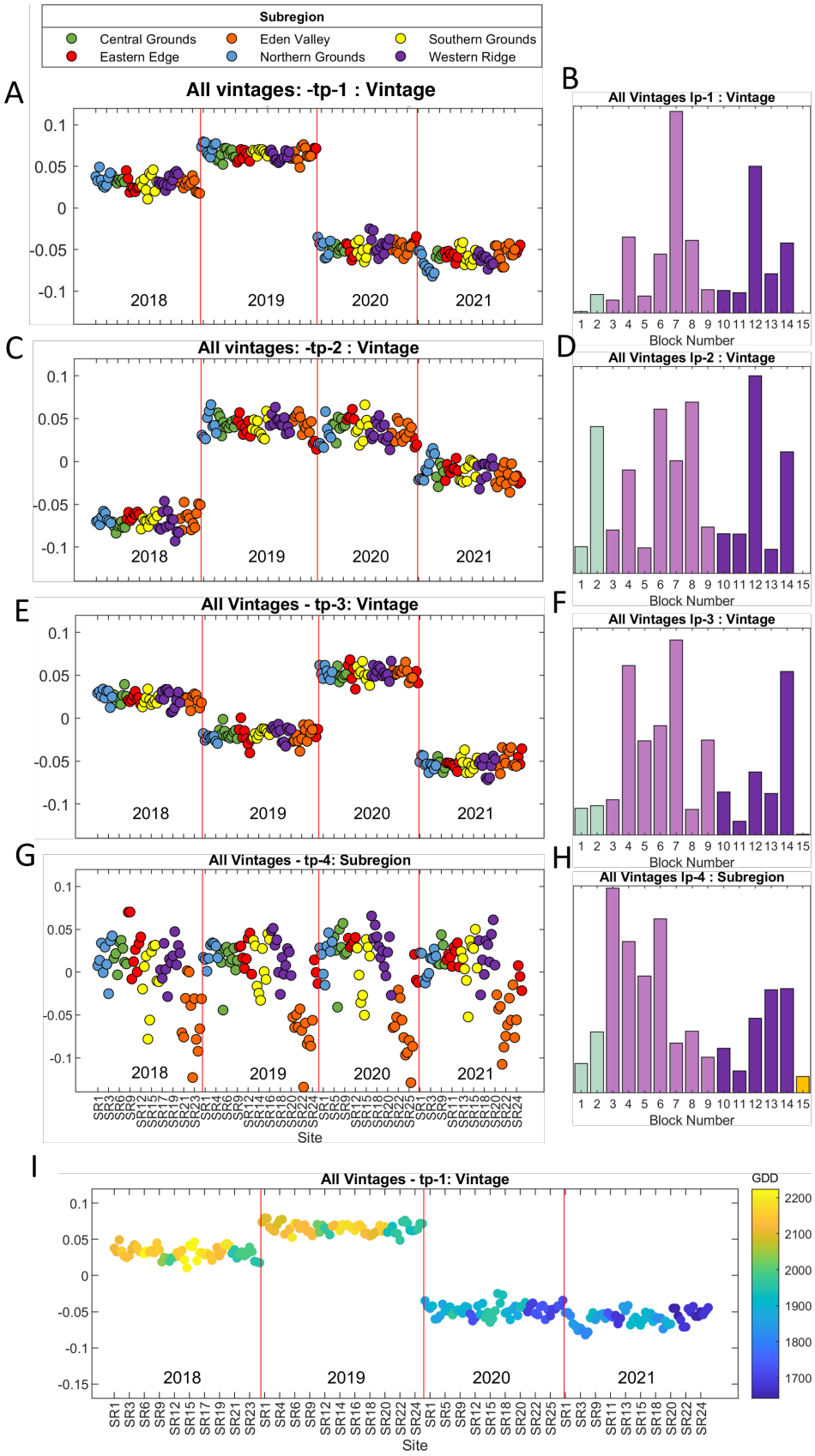


FIGURE 6. AMOPLS All Vintage (A, C, & E; tp1–3) and subregion (G; tp4) predictive scores from the "All Vintages" model for site (SR1–25) and colour-coded for subregion. Data block contributions to the corresponding predictive scores are shown in panels B, D, F, & H. The score for the dominant vintage level (tp1) is shown (I) with colour coding for the mean Growing Degree Days (GDD) of subregion sites for the experiment. Data block numbers correspond to Table 1.

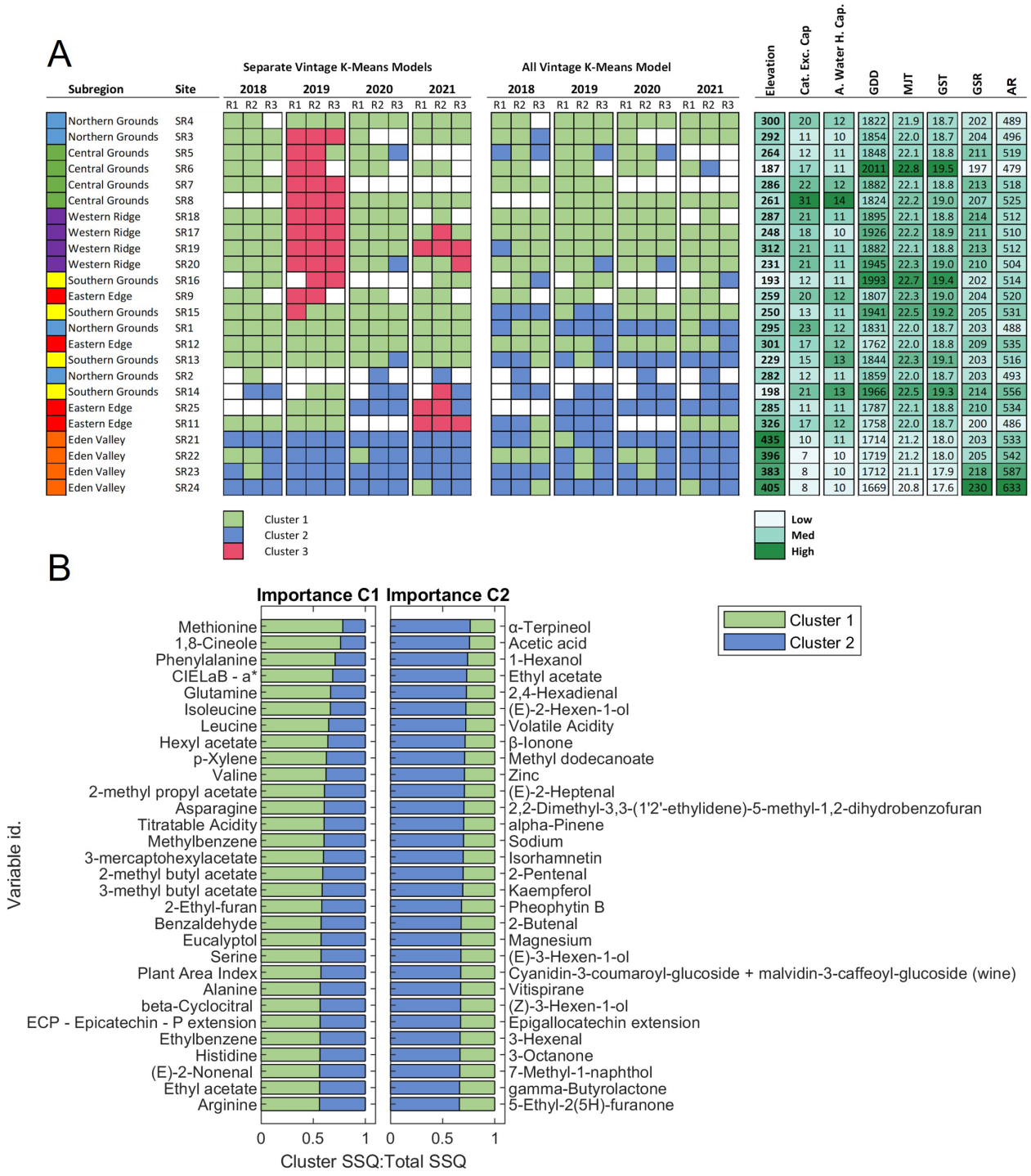


FIGURE 7. (A) *k*-means cluster solutions for sites and replicates for separate vintage models and the combined four vintage model. (B) The thirty most influential variables for sample assignment to cluster 1 and cluster 2 for the "All Vintages" model are shown with variable importance measured by the ratio of variable sum of squares for cluster samples to the total variable sum of squares in the dataset. Variable contribution to each cluster is illustrated in the alternative colour.

Each site replicate in each vintage is presented as a cluster identified by colour code for either cluster 1, 2, or 3. White space indicates missing data for that site replicate combination and exclusion from the model. Separate vintage *k*-means solutions are shown for each replicate and compared with the cluster assignment for the all-vintage model. Elevation = site m Australian height datum, Cat. Exc. Cap = profile weighted (5–60 cm) mean value for soil cation exchange capacity (cmol+/kg); A. Water H. Cap = profile weighted (5–60 cm) mean value for soil available water holding capacity (%); GDD = Season growing degree days (°C; base of 10 with no upper cap); MJT = Mean January Temperature (°C); GST = Mean growing season temperature (°C); GSR = Growing season rainfall (mm); AR = Annual rainfall (mm).

of grape composition, particularly amino acids that contribute to yeast assimilable nitrogen, which impact yeast formation of wine aroma compounds contributing to wine style. Viticultural management that increases the *de novo* supply of essential nitrogenous compounds for yeast biomass and fermentation is an important aspect of terroir.

4. Bagged classification tree ensemble – random forest

Bagged and boosted decision trees have become a standard classification approach for many machine learning problems in the literature. Typically, ensembles of decision trees are trained and tested using massive datasets that may range from several hundred thousand to billions of samples, with smaller numbers of features and computational speed is important (Chen & Guestrin, 2016). In the present investigation, the size of the dataset is modest and the ability to accurately predict *Subregion* without bias or overfitting is important. Random forest ensembles are reported to perform well for class prediction with complex datasets and they are quite robust to overfitting (Stavropoulos *et al.*, 2020). However, misclassification errors can increase when the proportion of class-relevant features in the dataset used for prediction is small (Friedman, 2017). A recursive elimination of low-ranked features improved RF performance (Figure 3A) with the selected model using 88 attributes with 100 grown trees. Model performance for the training and test sets is comparable based on AUROC (Figure 3B), indicating appropriate model fit. Good predictive acuity with positive predictive values (Figure 3C) for *Subregions* range from 60 % (Central Grounds) to 100 % (Eastern Edge). Samples from Eastern Edge, Southern, and Central Grounds subregions were predicted with the highest false negative rate (58 %, 33 %, and 33 %, respectively), and the positioning of these subregions in the MDS plot (Figure 3D) shows considerable overlap of confidence intervals. The two dimensions of the MDS plot account (Figure 3D) for only around 21 % of data variance, indicating the complexity of the dataset. These results infer that *Subregions* Eastern Edge, Southern, and Central Grounds are difficult to discern using the measured attributes, with Southern Ground samples frequently classed as either Eden Valley or Western Ridge; and Eastern Edge samples most frequently classed as Central Grounds or Southern Grounds. Samples from other subregions are well classified with high levels of prediction and low false negative results. The important attributes for the RF model (Figure 3E) are comprised of features from data blocks for grape volatiles, including diethyl succinate, a noted marker for wine ageing (Ubeda *et al.*, 2019), grape fatty acids, wine metals, wine volatiles, and vine performance, indicating subtle but important viticultural differences amongst the sites chosen in each subregion. Of interest is the high ranking of styrene (also in the BCT model), which is a measured attribute of grape composition. The presence of styrene is reported to be a marker for geographical origin for Molixiang table grapes (Feng *et al.*, 2022) and has also been reported to be present in Pinot blanc wines from Italy (Darnal *et al.*, 2024) and in a range of other foods (Steele *et al.*, 1994). The origins of styrene in both grapes and wine are thought to arise either through contamination by petroleum-based products and packaging (Ajaj *et al.*, 2021) or through biotransformation of trans-cinnamic acid (McKenna & Nielsen, 2011) and phenylalanine (Kim *et al.*, 2019). Reports of styrene as a marker for the geographical origin of grapes suggest this compound is deserving of closer attention in future studies.

5. Boosted classification trees – AdaBoost.M2

The BCT model's overall errors and performance (Figure 4A) were on par with the RF models for the *Subregion* test sample class prediction, with the selected model chosen after 18 iterations, which possessed 55 features. The AUROC curves (Figure 4B) show prediction outcomes for the independent test data set and the training sample sets have similar performance characteristics for each subregion, indicating an appropriately trained model without overfitting. The selected model performance was comparable to the RF results with high positive predictive value for *Subregions* Southern Grounds, Eden Valley, and Northern Grounds (90 %, 87 %, and 83 %, respectively) (Figure 4C). The Southern Grounds subregion also has a false negative rate of 25 % with samples predicted to belong to the Western Ridge or Eden Valley subregions. Consequently, the positive predictive value for the Western Ridge subregion is reduced to 73 % and false discovery rate to 27 %. The Eastern Edge subregion is modelled with a positive predictive value of 78 % and false negative rate of 42 % with samples misclassified to Western Ridge and Central Grounds. Notable are the high positive predictive values for the Southern Grounds and Eden Valley. Feature importance for the BCT model (Figure 4D) is dominated by a high value for diethyl succinate relative to the other features, followed by a range of grape volatile compounds, wine metals including sodium, strontium, cobalt and calcium, measures of vineyard performance and grape phenolics. Diethyl succinate may increase in concentration during wine ageing (Ubeda *et al.*, 2019) and is also reported to be elevated in wines made from grapes from nitrogen-deficient vineyards (Dienes-Nagy *et al.*, 2020) possessing low yeast assimilable nitrogen (Garde-Cerdán & Ancín-Azpilicueta, 2008; Ubeda *et al.*, 2019). Strontium and cobalt concentrations vary according to geographical origin for some agricultural commodities (Richter *et al.*, 2019), and isotope ratios (Reyrolle *et al.*, 2023) are increasingly useful for authentication.

6. Neural network

The neural network model using all measured features of the samples was the least successful classification model for the prediction of *Subregion* class in the training set, possibly due to the relatively small sample size, which is less suited to NN analysis and the complex architecture that makes the output difficult to interpret. The optimised neural network model ROC curves indicate the model is most robustly tuned to predict Central and Northern Grounds test samples, but lacks predictive accuracy for other subregion classes, as evidenced by the performance difference between the training and test sample sets (Figure 5A). The Northern Grounds class has the highest PPV and TPR at 64 and 70 %, respectively (Figure 5B), and Southern Grounds had the

highest FDR and FNR at 75 and 73 %, respectively. NNs are computationally intensive, with complex architectures that make it difficult to fully understand model outcomes. Recent gains in computation speed and the development of game theory approaches to determine the relative importance of features to NN outcomes have now facilitated a rise in their popularity (Mendez *et al.*, 2020). Shapely profiles of features of importance in the NN model (Figure 5C) demonstrate the complex interplay of wine metals (calcium, nickel, strontium, manganese and cobalt), grape volatiles and vine performance measures for subregion discrimination. Recent profiles of wine elemental profiles demonstrated their potential for confirmation of authenticity for wines of a limited geographical origin (Astray *et al.*, 2021) supporting our observations that wine elemental analysis may be important markers for subregional discrimination in combination with other measured attributes. There are no dominant features of importance in the optimised NN model, unlike the decision tree classification models, with the ranked features contributing to sample classification at a similar mean absolute Shapley value and differentially for class.

7. Limitations of this work

This four year investigation created a rich data set of measured attributes for vine performance, grape and wine composition, and wine sensory domains pertaining to Shiraz wines from the Barossa Zone GI, and adds to an existing comprehensive body of knowledge of wine styles and composition from this GI (Bonada *et al.*, 2021; Bramley & Ouzman, 2022; Johnson *et al.*, 2013; Kustos *et al.*, 2020; Li *et al.*, 2021; Moran *et al.*, 2019; Pearson *et al.*, 2020; Pearson *et al.*, 2021; Schmidtke *et al.*, 2024). Terroir studies are limited in experimental design as boundaries between geographical regions are rarely entirely solid or easily identified, with changes occurring within a gradient. Spatial autocorrelation, between sample sites arise, that is, plots close to one another will tend to be more similar than those further away, potentially limiting independence of samples within specific groups or replicates within vineyard sites. Sample sites within vineyards in the present investigation were chosen for maximum heterogeneity based upon standard geospatial measures. No *a priori* information for subregion is used in *k*-means clustering models, and an important question, therefore, arises: 'Is within vineyard variation greater, or smaller, than the variations between the subregional demarcations used in the study? Outcomes from *k*-means clustering suggest that, in some instances, variation between vineyard replicates is large, which may arise either from truly natural variations imparted from the site (Bramley *et al.*, 2011; Bramley *et al.*, 2017) or variations arising from methods used for sample analysis. An issue of sample independence becomes problematic for statistical tests of variation between two (or more) sample sets, and the test for significance is based upon continuous probability distributions defined by numerator and denominator degrees of freedom, such as the F-ratio. An increase in type 1 errors, *i.e.*, false positives, can occur in circumstances where inflated degrees of freedom arise through non-independence of samples. Unlike classical Analysis of Variance, AMOPLS makes use of permutation tests for determining empirical probability for differences between sample groups. The ensemble decision trees and artificial neural networks in the present investigation used observed (centred and scaled) data for classification. There is very limited literature on the impact of sample independence for machine learning outcomes, and we emphasise that decision trees and neural networks were originally developed using massive datasets of samples with a smaller number of features than in the present investigation. Additional limitations arise from the vineyard replicates, whilst chosen for maximum heterogeneity, may not be completely independent and therefore supervised discrimination algorithms may be overly optimistic in classification outcomes. Thus, another question arises: 'Can useful information be obtained through the application of these algorithms to the existing data? We have sought to be cautiously pragmatic in our interpretation of model outcomes. We do not infer that aspects of site, subregion or vineyard conditions cause specific compositional outcomes in all vineyards, but limit our discussion to the current samples and the Barossa Zone GI. Moreover, our focus is upon understanding how different data decomposition methods select features differentially, even when the same feature ranking approach is used, and, therefore, how researchers may differentially interpret experimental influence on sample composition, depending upon the choice of model algorithm.

Importantly, the current investigation demonstrates that different approaches to modelling terroir-related data emphasise different features within a dataset depending on the algorithmic approach used. As there are no universal methods applicable to ranking feature importance across different models, it is important to understand how differences between modelling outcomes arise. The AMOPLS approach is a linear model, and variable importance in projection (VIP) is a useful measure of feature contribution to predictive scores (González-Ruiz *et al.*, 2017). VIP is a cumulative measure of the weight of each variable relative to the others (Lu *et al.*, 2014), with the average of the squared VIP values equal to 1, which thereby establishes the criterion for feature importance (Akarachantachote *et al.*, 2014). An advantage of AMOPLS for understanding feature contributions to modelled factors is the removal of orthogonal information, *i.e.*, structured information that is uncorrelated to the factors of interest, from the data, and this facilitates interpretation of the variable loadings, but will not impact the VIP for each feature. *k*-Means also makes use of linear data combinations to find the distances between samples within the multivariate spaces associated with the data set, and in this investigation, feature importance has been determined from the ratio of feature sums of squares for in-class samples relative to the entire data set.

Biological systems are rarely linear, and the application of non-linear modelling, including ensembles of decision trees and neural networks, are increasingly popular for modelling grape and wine data (Hensel *et al.*, 2025; Sáenz-Navajas *et al.*, 2025). Understanding the contribution

of features to non-linear model performance is important as it helps to avoid biases from sample selection, or hidden artefacts that impact upon predictive acuity, are not inadvertently present in the dataset (Guidotti *et al.*, 2018). Common approaches for assessing feature importance in non-linear models are based upon permutation of features in the model and measuring performance loss (Altmann *et al.*, 2010); calculation of feature Gini impurity, which assess the likelihood of an incorrect sample class assignment based on the distribution of class values at the feature level (Raileanu & Stoffel, 2004), and more recently calculation of Shapley values (Štrumbelj & Kononenko, 2014). Shapley values attempt to quantify a feature contribution to model performance based on permutations of sequential feature addition at a point of interest (Merrick & Taly, 2020).

In classic Shapley calculations, models are evaluated using all permutations of feature additions to the model. With every possible combination of feature addition, computational requirements increase factorially, and given the computational requirements for calculating Shapley values with a large number of features or samples, varying approaches for estimating Shapley values have been devised (Chen *et al.*, 2022; Lundberg & Lee, 2017). Approaches for estimating Shapley values may vary, such as the selection of feature subsets, approximate feature Shapley values from model weights for linear models, or use back propagation for deep learning models (Lundberg & Lee, 2017). Importantly, limitations in the interpretation of Shapley values for features have been raised. The approaches to Shapley calculations (Merrick & Taly, 2020) and the order for feature assessment (Huang & Marques-Silva, 2024) can significantly influence outcomes. Moreover, Shapley values are not explanatory as they are a measure of the relative importance of features for a model prediction. Thus, some features with high Shapley scores may appear misleading in terms of model importance, and this is especially noted with increasing numbers of features (Huang & Marques-Silva, 2024). Nonetheless, Shapley values have gained widespread acceptance for explaining block box methods for prediction. In this study, we present non-linear model outcomes with measures of feature importance using Shapley values, and each model has somewhat different feature importance scores evident in Figures 3E, 4D, and 5C. An advantage of Shapley values in the present investigation is the demonstration of the non-linear contribution of features to subregional classification models, evident by the different absolute Shapley values coded for subregion. In the present investigation, vastly different outcomes for modelling were noted when feature importance was measured using permutation (RF) and Gini values (BCT) (results not shown). An emphasis on feature importance enables interpretation of the impact of terroir on measures of vine performance, grape and wine compositional measures. Known limitations for Shapley values for feature contribution to model outcomes demonstrate the importance of careful reflection, rather than placing blind trust in model outcomes.

CONCLUSION

Different approaches to modelling terroir data sets emphasise varying levels of feature importance, which in turn may influence the interpretation of experimental design factors associated with terroir studies. Linear and non-linear multivariate models exemplifying temporal and spatial data trends from a longitudinal viticultural and wine study have been used to characterise terroir influences at the within-region level. In terroir studies, the approach to data modelling is frequently determined through the researcher's experience with statistical methodology, with two broad approaches being classification and clustering techniques. Some researchers prefer clustering approaches because there are no predetermined data structures to accommodate within the model, thereby providing a level of legitimacy to the model's outcomes (Bramley & Ouzman, 2022). Classification models, on the other hand, employ the inherent knowledge of the experimental design to fit models to the sample class whilst providing information about feature importance. Both approaches have merit; however, all models can be overfitted, and care must be taken to ensure rigorous validation of models prior to interpretation, which is especially important for machine learning methods originally developed using massive data sets. In the present investigation, clustering and classification methods have been used to develop an understanding of terroir-associated factors for Shiraz grown in the Barossa GI for a complex data set of 15 data blocks of attributes measuring vine growth, grape and wine composition, including sensory appraisal of the resulting wines.

The approaches and selection of a statistical model to characterise the data are important considerations, as various methods exemplify and weight features differentially, which may influence the interpretation and importance of model outcomes by researchers undertaking the work. Researcher awareness of inherent algorithmic biases is therefore essential to ensure appropriate model interpretation with respect to determining terroir impact upon grape and wine composition. Irrespective of the chosen approach for data modelling, interesting information can be derived that relates subregion influences to measures of composition, but feature importance may not be consistent between the models.

FUNDING STATEMENT

This work was jointly funded by Charles Sturt University (CSU), Adelaide University (UA), The Australian Wine Research Institute (AWRI), The South Australian Research and Development Institute (SARDI), the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Wine Australia (project UA1602) whose contribution derived from levies collected from Australia's grapegrowers and winemakers with matching funds from the Australian Government. UA, AWRI, SARDI and CSIRO are members of the Wine Innovation Cluster, based at the Waite Campus, Adelaide.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge Barossa Australia and, in particular, the growers who allowed us to sample material from their properties and supplied information about their vineyards and management strategies. Many staff provided technical support for the collection and analysis of samples, including Ms Annette James, Mr Gaston Sepulveda, Ms Emily Nicholson, Ms Sue Maffei, Song Qi, Flynn Watson and Maddy Jiang. Dr Lira Souza-Gonzaga assisted in the collection of sensory data and additional sensory advice was provided by Dr I. Leigh Francis. The guidance to the program provided by Dr Paul Smith is also greatly appreciated. Sentek and GreenBrain supplied us with discounted soil moisture sensors and weather stations. The staff of the Spatial Data Analysis Network at Charles Sturt University are gratefully acknowledged for assistance in creating figures.

REFERENCES

Abdi, H., Williams, L. J., Valentin, D., & Bennani-Dosse, M. (2012). STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*, 124-167. https://doi.org/10.1002/wics.198

Aggarwal, C. C. (2023). *Neural Networks and Deep Learning : A Textbook*. Springer International Publishing AG. http://ebookcentral.proquest.com/lib/csuau/detail.action?docID=30620507. https://doi.org/10.1007/978-3-031-29642-0

Aggarwal, C.C., Hinneburg, A., Keim, D.A. (2001). On the Surprising Behavior of Distance Metrics in High Dimensional Space. In: Van den Bussche, J., Vianu, V. (eds) Database Theory — ICDT 2001. ICDT 2001. Lecture Notes in Computer Science, vol 1973. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-44503-X_27

Ajaj, A., J'Bari, S., Ononogbo, A., Buonocore, F., Bear, J. C., Mayes, A. G., & Morgan, H. (2021). An Insight into the Growing Concerns of Styrene Monomer and Poly(Styrene) Fragment Migration into Food and Drink Simulants from Poly(Styrene) Packaging. *Foods*, *10*(5). https://doi.org/10.3390/foods10051136

Akarachantachote, N., Chadcham, S., & Saithanu, K. (2014). Cutoff Threshold of Variable Importance in Projection for Variable Selection. *International Journal of Pure and Applied Mathematics*, *94*(3), 307-322. https://doi.org/10.12732/ijpam.v94i3.2

Alem, H., Rigou, P., Schneider, R., Ojeda, H., & Torregrosa, L. (2019). Impact of agronomic practices on grape aroma composition: a review. *Journal of the Science of Food and Agriculture*, *99*(3), 975-985. https://doi.org/10.1002/jsfa.9327

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, *26*(10), 1340-1347. https://doi.org/10.1093/bioinformatics/btq134

Antalick, G., Šuklje, K., Blackman, J. W., Meeks, C., Deloire, A., & Schmidtke, L. M. (2015). Influence of grape composition on red wine ester profile: comparison between Cabernet Sauvignon and Shiraz cultivars from Australian warm climate. *Journal of Agricultural and Food Chemistry*, *63*(18), 4664-4672. https://doi.org/10.1021/acs.jafc.5b00966

Antalick, G., Šuklje, K., Blackman, J. W., Schmidtke, L. M., & Deloire, A. (2021). Performing sequential harvests based on berry sugar accumulation (mg/berry) to obtain specific wine sensory profiles. *OENO One*, *55*(2), 131-146. https://doi.org/10.20870/oeno-one.2021.55.2.4527

Astray, G., Martinez-Castillo, C., Mejuto, J.-C., & Simal-Gandara, J. (2021). Metal and metalloid profile as a fingerprint for traceability of wines under any Galician protected designation of origin. *Journal of Food Composition and Analysis*, *102*, 104043. https://doi.org/10.1016/j.jfca.2021.104043

Ballester, J. (2021). *In search of the taste of terroir - a challenge for sensory science.* XIIIth International Terroir Adelaide. https://ives-openscience.eu/6673

Boccard, J., & Rudaz, S. (2016). Exploring omics data from designed experiments using analysis of variance multiblock orthogonal partial least squares. *Analytica Chimica Acta*, *920*, 18-28. https://doi.org/10.1016/j.aca.2016.03.042

Bonada, M., Catania, A. A., Gambetta, J. M., & Petrie, P. R. (2021). Soil water availability during spring modulates canopy growth and impacts the chemical and sensory composition of Shiraz fruit and wine. *Australian Journal of Grape and Wine Research*, *27*(4), 491-507. https://doi.org/10.1111/ajgw.12506

Bramley, R. G. V., & Ouzman, J. (2022). Underpinning terroir with data: on what grounds might subregionalisation of the Barossa zone geographical indication be justified? *Australian Journal of Grape and Wine Research*, *28*(2), 196-207. https://doi.org/10.1111/ajgw.12513

Bramley, R. G. V., Ouzman, J., & Boss, P. K. (2011). Variation in vine vigour, grape yield and vineyard soils and topography as indicators of variation in the chemical composition of grapes, wine and wine sensory attributes. *Australian Journal of Grape and Wine Research*, *17*(2), 217-229. https://doi.org/10.1111/j.1755-0238.2011.00136.x

Bramley, R. G. V., Siebert, T. E., Herderich, M. J., & Krstic, M. P. (2017). Patterns of within-vineyard spatial variation in the 'pepper' compound rotundone are temporally stable from year to year. *Australian Journal of Grape and Wine Research*, *23*(1), 42-47. https://doi.org/10.1111/ajgw.12245

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Brillante, L., Bonfante, A., Bramley, R. G. V., Tardaguila, J., & Priori, S. (2020). Unbiased Scientific Approaches to the Study of Terroir Are Needed! [Opinion]. *Frontiers in Earth Science*, *8*. https://doi.org/10.3389/feart.2020.539377

Bunde, A. (2023). The Different Types of Noise and How They Effect Data Analysis. *Chemie Ingenieur Technik*, *95*(11), 1758-1767. https://doi.org/https://doi.org/10.1002/cite.202300031

Ceccarelli, G., Grandi, A., & Magagnoli, S. (2010). The "Taste" of typicality. *Food and History*, *8*(2), 45-76. https://doi.org/10.1484/J.FOOD.1.102217

Chen, H., Lundberg, S. M., & Lee, S.-I. (2022). Explaining a series of models by propagating Shapley values. *Nature Communications*, *13*(1), 4512. https://doi.org/10.1038/s41467-022-31384-3

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system* Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA. https://doi.org/10.1145/2939672.2939785

Cosme, F., Milheiro, J., Pires, J., Guerra-Gomes, F. I., Filipe-Ribeiro, L., & Nunes, F. M. (2021). Authentication of Douro DO monovarietal red wines based on anthocyanin profile: Comparison of partial least squares – discriminant analysis, decision trees and artificial neural networks. *Food Control*, *125*, 107979. https://doi.org/https://doi.org/10.1016/j.foodcont.2021.107979

Darnal, A., Poggesi, S., Merkytė, V., Longo, E., & Boselli, E. (2024). South-Tyrolean pinot blanc identity: Exploration of chemical and sensory profile changes ascribed to vineyard locations and winemaking variables. *Food Chemistry: X*, *24*, 101824. https://doi.org/10.1016/j.fochx.2024.101824

Dienes-Nagy, Á., Marti, G., Breant, L., Lorenzini, F., Fuchsmann, P., Baumgartner, D., Zufferey, V., Spring, J.-L., Gindro, K., Viret, O., Wolfender, J.-L., & Rösti, J. (2020). Identification of putative chemical markers in white wine (Chasselas) related to nitrogen deficiencies in vineyards. *OENO One*, *54*(3), 583-599. https://doi.org/10.20870/oeno-one.2020.54.3.3285

Drappier, J., Thibon, C., Rabot, A., & Geny-Denis, L. (2019). Relationship between wine composition and temperature: Impact on Bordeaux wine typicity in the context of global warming—Review. *Critical Reviews in Food Science and Nutrition*, *59*(1), 14-30. https://doi.org/10.1080/10408398.2017.1355776

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802-813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Fairbairn, S., McKinnon, A., Musarurwa, H. T., Ferreira, A. C., & Bauer, F. F. (2017). The impact of single amino acids on growth and volatile aroma production by *Saccharomyces cerevisiae s*trains [Original Research]. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.02554

Feng, T., Sun, J., Song, S., Wang, H., Yao, L., Sun, M., Wang, K., & Chen, D. (2022). Geographical differentiation of Molixiang table grapes grown in China based on volatile compounds analysis by HS-GC-IMS coupled with PCA and sensory evaluation of the grapes. *Food Chemistry: X*, *15*, 100423. https://doi.org/10.1016/j.fochx.2022.100423

Friedman, J. (2017). The elements of statistical learning data mining, inference, and prediction (Second ed.). Springer. https://doi.org/https://doi.org/10.1007/978-0-387-84858-7

Garde-Cerdán, T., & Ancín-Azpilicueta, C. (2008). Effect of the addition of different quantities of amino acids to nitrogen-deficient must on the formation of esters, alcohols, and acids during wine alcoholic fermentation. *LWT - Food Science and Technology*, *41*(3), 501-510. https://doi.org/https://doi.org/10.1016/j.lwt.2007.03.018

González-Ruiz, V., Pezzatti, J., Roux, A., Stoppini, L., Boccard, J., & Rudaz, S. (2017). Unravelling the effects of multiple experimental factors in metabolomics, analysis of human neural cells with hydrophilic interaction liquid chromatography hyphenated to high resolution mass spectrometry. *Journal of Chromatography A*, *1527*, 53-60, Article 174. https://doi.org/10.1016/j.chroma.2017.10.055

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, *51*(5), Article 93. https://doi.org/10.1145/3236009

Hensel, M., Vestner, J., Fahrer, J., & Durner, D. (2025). Evaluation of Machine Learning Algorithms to Classify Blanc de noir Wines with Spectrophotometric Data. *American Journal of Enology and Viticulture*, *76*(1), 0760006. https://doi.org/10.5344/ajev.2024.24029

Hinwood, A. L., Berko, H. N., Farrar, D., Galbally, I. E., & Weeks, I. A. (2006). Volatile organic compounds in selected micro-environments. *Chemosphere*, *63*(3), 421-429. https://doi.org/https://doi.org/10.1016/j.chemosphere.2005.08.038

Huang, X., & Marques-Silva, J. (2024). On the failings of Shapley values for explainability. *International Journal of Approximate Reasoning*, *171*, 109112. https://doi.org/10.1016/j.ijar.2023.109112

Johnson, T. E., Hasted, A., Ristic, R., & Bastian, S. E. P. (2013). Multidimensional scaling (MDS), cluster and descriptive analyses provide preliminary insights into Australian Shiraz wine regional characteristics. *Food Quality and Preference*, *29*(2), 174-185. https://doi.org/10.1016/j.foodqual.2013.03.010

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). *LightGBM: A highly efficient gradient boosting decision tree.* https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

Kim, H. W., Lee, S. M., Seo, J.-A., & Kim, Y.-S. (2019). Effects of pH and Cultivation Time on the Formation of Styrene and Volatile Compounds by Penicillium expansum. *Molecules*, *24*(7). https://doi.org/10.3390/molecules24071333

Kustos, M., Gambetta, J. M., Jeffery, D. W., Heymann, H., Goodman, S., & Bastian, S. E. P. (2020). A matter of place: Sensory and chemical characterisation of fine Australian Chardonnay and Shiraz wines of provenance. *Food Research International*, *130*, 108903. https://doi.org/10.1016/j.foodres.2019.108903

Li, J., Qian, J., Chen, J., Ruiz-Garcia, L., Dong, C., Chen, Q., Liu, Z., Xiao, P., & Zhao, Z. (2025). Recent advances of machine learning in the geographical origin traceability of food and agro-products: A review. *Comprehensive Reviews in Food Science and Food Safety*, *24*(1), e70082. https://doi.org/https://doi.org/10.1111/1541-4337.70082

Li, S., Blackman, J. W., & Schmidtke, L. M. (2021). Exploring the regional typicality of Australian Shiraz wines using untargeted metabolomics. *Australian Journal of Grape and Wine Research*, *27*(3), 378-391. https://doi.org/10.1111/ajgw.12493

Lu, B., Castillo, I., Chiang, L., & Edgar, T. F. (2014). Industrial PLS model variable selection using moving window variable importance in projection. *Chemometrics and Intelligent Laboratory Systems*, *135*, 90-109. https://doi.org/10.1016/j.chemolab.2014.03.020

Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions* Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA.

Ma, P., Zhang, Z., Jia, X., Peng, X., Zhang, Z., Tarwa, K., Wei, C.-I., Liu, F., & Wang, Q. (2024). Neural network in food analytics. *Critical Reviews in Food Science and Nutrition*, *64*(13), 4059-4077. https://doi.org/10.1080/10408398.2022.2139217

Manicka, S., Johnson, K., Levin, M., & Murrugarra, D. (2023). The nonlinearity of regulation in biological networks. *npj Systems Biology and Applications*, *9*(1), 10. https://doi.org/10.1038/s41540-023-00273-w

Marini, F. (2020). 3.24 - Non-linear Modeling: Neural Networks☆. In S. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive Chemometrics (Second Edition)* (pp. 519-541). Elsevier. https://doi.org/10.1016/B978-0-12-409547-2.14893-0

McKenna, R., & Nielsen, D. R. (2011). Styrene biosynthesis from glucose by engineered E. coli. *Metabolic Engineering*, *13*(5), 544-554. https://doi.org/10.1016/j.ymben.2011.06.005

Mendez, K. M., Broadhurst, D. I., & Reinke, S. N. (2020). Migrating from partial least squares discriminant analysis to artificial neural networks: a comparison of functionally equivalent visualisation and feature contribution tools using jupyter notebooks. *Metabolomics*, *16*(2), 17. https://doi.org/10.1007/s11306-020-1640-0

Merrick, L., Taly, A. (2020). The Explanation Game: Explaining Machine Learning Models Using Shapley Values. In: Holzinger, A., Kieseberg, P., Tjoa, A., Weippl, E. (eds) Machine Learning and Knowledge Extraction. CD-MAKE 2020. Lecture Notes in Computer Science(), vol 12279. Springer, Cham. https://doi.org/10.1007/978-3-030-57321-8_2

Moran, M., Petrie, P., & Sadras, V. (2019). Effects of Late Pruning and Elevated Temperature on Phenology, Yield Components, and

Berry Traits in Shiraz. *American Journal of Enology and Viticulture*, *70*(1), 9. https://doi.org/10.5344/ajev.2018.18031

Oliveira, R. S., Costa, L. S., de Almeida Santiago, H., da Silva Mutz, Y., de Oliveira Faria, R., Figueiredo, L. P., Guimarães, P. H. S., Curi, N., & de Menezes, M. D. (2025). Decoding local terroir: Data mining to predict sensory profiles of coffee beverage. *Agricultural Systems*, *230*, 104487. https://doi.org/https://doi.org/10.1016/j.agsy.2025.104487

Pearson, W., Schmidtke, L., Francis, I., Carr, B., & Blackman, J. (2020). Characterising inter-and intra-regional variation in sensory profiles of Australian Shiraz wines from six regions. *Australian Journal of Grape and Wine Research*, *26*(4), 372-384. https://doi.org/10.1111/ajgw.12455

Pearson, W., Schmidtke, L., Francis, I. L., Li, S., Hall, A., & Blackman, J. (2021). Regionality in Australian Shiraz: compositional and climate measures that relate to key sensory attributes. *Australian Journal of Grape and Wine Research*. https://doi.org/10.1111/ajgw.12499

Raileanu, L. E., & Stoffel, K. (2004). Theoretical Comparison between the Gini Index and Information Gain Criteria. *Annals of Mathematics and Artificial Intelligence*, *41*(1), 77-93. https://doi.org/10.1023/B:AMAI.0000018580.96245.c6

Reyrolle, M., Bareille, G., Epova, E. N., Barre, J., Bérail, S., Pigot, T., Desauziers, V., Gautier, L., & Le Bechec, M. (2023). Authenticating teas using multielement signatures, strontium isotope ratios, and volatile compound profiling. *Food Chemistry*, *423*, 136271. https://doi.org/10.1016/j.foodchem.2023.136271

Richter, B., Gurk, S., Wagner, D., Bockmayr, M., & Fischer, M. (2019). Food authentication: Multi-elemental analysis of white asparagus for provenance discrimination. *Food Chemistry*, *286*, 475-482. https://doi.org/10.1016/j.foodchem.2019.01.105

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

Sáenz-Navajas, M.-P., Ferreira, C., Bastian, S. E. P., & Jeffery, D. W. (2025). Bagging and boosting machine learning algorithms for modelling sensory perception from simple chemical variables: Wine mouthfeel as a case study. *Food Quality and Preference*, *129*, 105494. https://doi.org/10.1016/j.foodqual.2025.105494

Sanjuán-Herráez, D., de la Osa, S., Pastor, A., & de la Guardia, M. (2014). Air monitoring of selected volatile organic compounds in wineries using passive sampling and headspace-gas chromatography–mass spectrometry. *Microchemical Journal*, *114*, 42-47. https://doi.org/https://doi.org/10.1016/j.microc.2013.11.017

Santo, S. D., Zenoni, S., Sandri, M., Lorenzis, G. D., Magris, G., Paoli, E. D., Gaspero, G. D., Fabbro, C. D., Morgante, M., Brancadoro, L., Grossi, D., Fasoli, M., Zuccolotto, P., Tornielli, G. B., & Pezzotti, M. (2018). Grapevine field experiments reveal the contribution of genotype, the influence of environment and the effect of their interaction (G×E) on the berry transcriptome. *The Plant Journal*, *93*(6), 1143-1159. https://doi.org/10.1111/tpj.13834

Schmidtke, L. M., Antalick, G., Suklje, K., Blackman, J. W., Boccard, J., & Deloire, A. (2020). Cultivar, site or harvest date: the gordian knot of wine terroir. *Metabolomics*, *16*(5), 52. https://doi.org/10.1007/s11306-020-01673-3

Schmidtke, L. M., Bastian, S. E. P., Bindon, K., Bonada, M., Boss, P. K., Bramley, R. G. V., Danner, L., Petrie, P. R., Gonzaga, L. S., & Collins, C. (2024). Exploring interactions between vineyard performance, grape and wine composition and subregional boundaries - the terroir of Barossa Shiraz. *Australian Journal of Grape and Wine Research*(1), 2622516. https://doi.org/10.1155/ajgw/2622516

Schüttler, A. (2013). *Influencing factors on aromatic typicality of wines from Vitis vinifera L. cv. Riesling - sensory, chemical and viticultural insights* - (Publication Number 2019) University of Bordeaux 2 and University of Giessen].

Stavropoulos, G., van Voorstenbosch, R., van Schooten, F.-J., & Smolinska, A. (2020). 3.32 - Random forest and ensemble methods. In S. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive Chemometrics (Second Edition)* (pp. 661-672). Elsevier. https://doi.org/10.1016/B978-0-12-409547-2.14589-5

Steele, D. H., Thornburg, M. J., Stanley, J. S., Miller, R. R., Brooke, R., Cushman, J. R., & Cruzan, G. (1994). Determination of styrene in selected foods. *Journal of Agricultural and Food Chemistry*, *42*(8), 1661-1665. https://doi.org/10.1021/jf00044a015

Štrumbelj, E., & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, *41*(3), 647-665. https://doi.org/10.1007/s10115-013-0679-x

Sumby, K. M., Grbin, P. R., & Jiranek, V. (2010). Microbial modulation of aromatic esters in wine: Current knowledge and future prospects. *Food Chemistry*, *121*(1), 1-16. https://doi.org/10.1016/j.foodchem.2009.12.004

Ubeda, C., Kania-Zelada, I., del Barrio-Galán, R., Medel-Marabolí, M., Gil, M., & Peña-Neira, Á. (2019). Study of the changes in volatile compounds, aroma and sensory attributes during the production process of sparkling wine by traditional method. *Food Research International*, *119*, 554-563. https://doi.org/10.1016/j.foodres.2018.10.032

van Leeuwen, C., Barbe, J.-C., Darriet, P., Geffroy, O., Gomès, E., Guillaumie, S., Helwi, P., Laboyrie, J., Lytra, G., Le Menn, N., Marchand, S., Picard, M., Pons, A., Schüttler, A., & Thibon, C. (2020). Recent advancements in understanding the terroir effect on aromas in grapes and wine. *OENO One*, *54*(4), 985-1006. https://doi.org/10.20870/oeno-one.2020.54.4.3983

van Leeuwen, C., Friant, P., Choné, X., Tregoat, O., Koundouras, S., & Dubourdieu, D. (2004). Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and Viticulture*, *55*(3), 207. https://doi.org/10.5344/ajev.2004.55.3.207

Wang, H., Liang, Q., Hancock, J. T., & Khoshgoftaar, T. M. (2024). Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. *Journal of Big Data*, *11*(1), 44. https://doi.org/10.1186/s40537-024-00905-w

Wine Australia. (2023). *Geographical indications*. Wine Australia. Retrieved 14/04/2023 from https://www.wineaustralia.com/labelling/register-of-protected-gis-and-other-terms/geographical-indications

Yu, B., Yuan, Z., Yu, Z., & Xue-song, F. (2022). BTEX in the environment: An update on sources, fate, distribution, pretreatment, analysis, and removal techniques. *Chemical Engineering Journal*, *435*, 134825. https://doi.org/https://doi.org/10.1016/j.cej.2022.134825

Zhu, J., Zou, H., Rosset, S., & Hastie, T. (2009). Multi-class AdaBoost. *Statistics and its Interface*, *2*, 349-360. https://doi.org/10.4310/SII.2009.v2.n3.a8