Towards Precise Grape Management Based on Multispectral and Hyperspectral Remote Sensing

A thesis presented in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in

Agriculture and Horticulture

at Massey University, Manawatū, New Zealand

Hongyi Lyu

2024

Abstract

Non-destructive and rapidly measurements of grape quality allow vine growers to make effective decisions for selective harvesting and vineyard management, which results in more economic and sustainable wine production. The objective of this dissertation was to investigate the potential of using non-destructive techniques at pre- and post-harvest stage to predict 'Pinot Noir' grape quality parameters (total soluble solids, TSS; and titratable acidity, TA). Three non-destructive techniques unmanned aerial vehicle (UAV) based on multispectral camera, visible-near infrared-shortwave infrared (VNIR-SWIR) spectroscopy, and hyperspectral imaging (HSI) were applied on commercial vineyards. In the multispectral study, the spectral reflectance measurements of vine canopy were acquired using an UAV during flowering stage. The acquired spectral data were combined with environmental ancillary data to predict grape TSS. UAV combined with ancillary data showed a moderate prediction performance, indicating grape growers can manage large vineyard areas efficiently, reducing labor and operational costs based on this technique. In the hyperspectral study, the spectral reflectance measurements of grape berries were acquired using VNIR-SWIR spectroscopy and HSI during pre- and post-harvest stages. VNIR-SWIR spectroscopy was used directly in the vineyard condition with a contact probe to ensure stable light conditions during pre-harvest stage. The acquired spectra data were pre-processed by various methods including principal component analysis and spectra pre-processing. The ability of using different spectral regions to predict grape TSS was validated with an external dataset. The spectroscopy under VNIR spectrum is sufficiency for accurate prediction, suggesting cost-effectiveness over VNIR-SWIR. HSI was used under laboratory conditions during post-harvest stage to predict and classify grape quality parameters (TSS and TA). The high level of classification accuracy achieved indicating HSI was adequate for grape TSS sorting purposes. However, classification of TA remained unsuccessful for industry application. Further study should consider to using different sensors to determine grape TA. This study demonstrated the ability of using different non-destructive techniques to predict grape quality parameters in 'Pinot Noir' grape variety, which would allow more accurate

and real-time assessment of grape quality, aiding in the optimal timing of harvest to enhance wine quality.

Acknowledgement

I would like to express my deepest gratitude to my main supervisor, Dr. Miles Grafton, whose continuous support, and encouragement have been invaluable throughout this study. His expertise and patience have been essential in guiding me to think as an independent researcher.

My sincere thanks also go to my co-supervisors, Associate Professor Thiagarajah Ramilan and Matthew Irwin. Associate Professor Ramilan taught me the fundamentals of the R programming language, which was crucial for my data analysis. Matthew Irwin provided invaluable instruction in the fundamentals of remote sensing and assisted me significantly with my fieldwork.

This project would not have been possible without the professional support of Eduardo Sandoval, whose expertise in multispectral and hyperspectral imagery operations was essential. His assistance and dedication greatly contributed to the success of this research.

Special thanks go to the staffs in Post Harvest Technology Lab and Associate Professor Gabor Kereszturi for providing the equipment used for grape quality analysis. Their support and resources were crucial for the completion of this research.

I would like to acknowledge the financial support provided by Professor Ren Jizhou Scholarship; Johannes August Anderson Postgraduate Scholarship; Turners and Growers Research Grant; Helen E Akers Postgraduate Scholarship; Mercury NZ Renewable Energy Research Bursary and Horticentre Trust Grant, which made this research possible.

I owe a special debt of gratitude to my parents, for their unconditional love and financial support during my study.

Table of Contents

List of Figures

List of Tables

Acronyms

1D	first derivative
3D	three-dimensional
3D-CNN	three-dimensional convolutional neural network
Abs	absorbance
Acc	accuracy
ANN	artificial neural network
AUD	australian dollar
AUC	specificity and area under the curve
BIP	band-interleaved-by-pixel
BIL	band-interleaved-by-line
BSQ	band sequential
BBCH	biologische bundesanstalt, bundessortenamt und cHemische industrie
Ce	classification error rate
CCC	Lin's concordance correlation coefficient
CN	cluston
CNNs	convolutional neural networks
DEM	digital elevation model
DOY	day of year
ES	stack ensemble learning
EC_a	electrical conductivity
FP	false positives
GAN	generative adversarial network
GIS	geographical information system
GPS	global position systems
HN	hua nui
HSI	hyperspectral imaging system
IQR	interquartile range

KNN	k-nearest neighbors
LDA	linear discriminant analysis
LSTM	long short-term memory neural network
Modified E-L	modified eichhorn and lorenz system
MLR	multivariate linear regression
MLP	multilayer perceptron
MSC	multiplicative scatter correction
NDSI	normalization different spectral indices
NDVI	normalized vegetation index
NIR	near infrared
nRMSE%	normalized root mean squared error percentage
PCA	principal component analysis
PCD	plant cell density
PCR	principal component regression
PCs	principal components
PFI	permutation feature importance
PLS-DA	partial least squares discriminant analysis
PLSR	partial least squares regression
PN	pencarrow
PV	precision viticulture
ReLU	rectified linear unit
RFR	random forest regression
RGB	red-green-blue
R^2	coefficient of determinations
RMSE	root mean squared error
ROI	region of interest
RPD	ratio prediction to deviation
RS	remote sensing

RSS	residual sum of squares
RTK	real-time kinematic
RNN	recurrent neural network
RW	raw reflectance
SG	savitzky–golay
SSRN	spatial residual network
SWP	stem water potential
SNV	standard normal variate
SVR	support vector regression
SVM	support vector machine
TA	titratable acidity
TC	trunk circumference
TSS	total soluble solids
UAV	unmanned aerial vehicle
VIs	vegetation indices
VIP	variable importance projection
VNIR-SWIR	visible-near infrared-shortwave infrared
WGAN	Wasserstain Generative Adversarial Network
XGBoost	extreme gradient boosting

1 Introduction

1.1 General Background

Wine grapes are a major horticultural crop in New Zealand, with around 41,860 hectares of land being used for wine grape production in 2023 (New Zealand Winegrowers, 2023). In New Zealand, ‘Sauvignon Blanc’ accounts for 65% of the total production area, and ‘Pinot Noir’ next at 14%. The New Zealand wine industry managed to achieve record exports totaling NZ$2,405 million in 2023, up 23.1% on the previous year (New Zealand Winegrowers, 2023). New Zealand has 11 wine producing regions across the two main islands with a range of mini-climates and soil types (Figure 1.1). ‘Pinot Noir’ is New Zealand’s most commonly planted red variety and second-most exported wine. Wairarapa and Central Otago are famous ‘Pinot Noir’ producing regions (Figure 1.1). Martinborough, situated within the Wairarapa region, stands out with its highly acclaimed ‘Pinot Noir’. Martinborough has a similar climate and soil profile with Burgundy which is the paragon for ‘Pinot Noir’.

Figure 1.1: The major vineyard regions in New Zealand (Winegrowers, 2023)

‘Pinot Noir’ is a notoriously fickle grape, as it’s sensitive to various environmental factors (Shaw, 2012). The tight berry clusters make them vulnerable to diseases like botrytis and mildew, as well as quality issues such as uneven ripening. Their thin skins also make them susceptible to damage from weather conditions such as excessive sun exposure. Compared with other wine grape varieties, the growing season of ‘Pinot Noir’ is relatively short. Its window of ripeness is narrow, often lasting just a few days (Nicholas, 2015). Thus, it is important to determine the optimal time to harvest the ‘Pinot Noir’ grapes when they are ripe. Harvesting too early can result in underripe flavors, whilst waiting too long can lead to overripe grapes, affecting the quality and value of the resulting wine (Grainger and Tattersall, 2016). Winemakers must closely monitor the grapes and be ready to harvest swiftly when they reach peak ripeness.

To date, the traditional method of determining grape maturity and quality involves analyzing total soluble solids, titratable acidity and pH based on destructive techniques. The process typically begins around veraison, the stage in the grapevine growth cycle when the berries start to change color and soften, marking the onset of ripening. From veraison, grape growers will take weekly samples of berries from different parts of the vineyard. As harvest approaches and the grapes near their peak ripeness, the sampling frequency may increase too daily. During each sampling day, grape growers pick around 200 berries per block and per variety. The maturity and quality levels of these 200 berries are used to represent the entire variety. However, many studies have shown that grape quality parameters varied spatially and temporally within blocks (Bramley, 2005; Baluja et al., 2013). Uniformity in grape berries is crucial for winemakers in achieving consistent quality in their wines. When grapes are uniform in terms of ripeness and quality, winemakers can expect a more predictable outcome throughout the winemaking process. This uniformity allows for better control over factors such as extraction, fermentation, and aging. Bramley et al., (2003) shows that selective harvesting, where grapes of different qualities are harvested and treated separately, can bring greater benefits to the winery. Using traditional destructive testing methods only evaluates the quality of a fraction of the grapes in a vineyard. The development of real-time non-destructive methods is preferable as it can monitor the maturity and quality levels of an entire vineyard.

1.2 Research Objectives

This study aims to explore the possibility of using various non-destructive methods applied during vine growth stages to predict ‘Pinot Noir’ quality parameters. Remote sensing imagery, VNIR-SWIR spectroscopy, and HSI system are explored as potential techniques. Remote sensing imagery is used to monitor the vine growth status from the canopy level. This information is related to vine vigor, water and nutrient status which affects the quality of wine grapes (Moghimi et al., 2020; Wei et al., 2022; Ferro et al., 2023). VNIR-SWIR spectroscopy and HSI are used to monitor the grape berry directly. Both VNIR-SWIR spectroscopy and HSI rely on detecting the reflection, absorption, or transmission of electromagnetic radiation on or within the surface of grape berries to assess their quality (Li et al., 2019).

1.3 Research Outline

This thesis is structured into seven chapters. The first two chapters cover the introduction and literature review. Following these are four stand-alone chapters, each developed as an individual paper submitted to peer-reviewed journals and cited accordingly in the author's publications list. Each of these chapters includes its own introduction, materials and methods, results, and discussion. The final chapter provides a research discussion and summary.

Chapter 1. It provides the study background about the New Zealand wine industry and outlines the objectives and thesis structure.

Chapter 2. It provides a literature review on current knowledge on winegrape, including phenology, quality attributes and precision viticulture It also introduces the principle of VNIR-SWIR and HSI, the application mode, data collection and analysis for quality determination of winegrapes.

Chapter 3. Investigates the ability of using remote sensing imagery and ancillary data to predict grape TSS. For this, an unmanned aerial vehicle (UAV) was used to acquire multispectral imagery of the vine growth status. Additionally other environmental variables from different sources were used to improve the model

performance of UAV imagery for predicting grape quality. This chapter also explores the different machine learning algorithms prediction performance.

Chapter 4. Investigates the feasibility of using proximal VNIR-SWIR sensors to assess grape TSS directly in the vineyard. In addition, this chapter also explores the comparative effectiveness of VNIR, SWIR, and VNIR-SWIR spectral data in predicting grape TSS under unstable environmental conditions. A large in-field multi-season spectra database (n = 1830) over two years (2023-2024) from three 'Pinot Noir' commercial vineyards were selected to develop different machine learning pipelines.

Chapter 5. Describes the ability of using hyperspectral imaging system to predict grape TSS and TA during the post-harvest stage. In this chapter, both, regression and classification models were adopted to determine grape quality parameters. This chapter also identifies the important wavelengths for explaining grape quality parameters.

Chapter 6. Uncovers the feasibility of using deep learning algorithms to generate synthetic grape hyperspectral reflectance data and improve the grape quality classification performance. For this, a Generative Adversarial Network was used to generate synthetic hyperspectral reflectance data. Various convolutional neural networks are used to analyze the hyperspectral reflectance data.

Chapter 7. It provides overall findings and discussion for this study and suggestions and perspectives for future research.

2 Literature Review

2.1 New Zealand Wine Grape

2.1.1 General Characteristics

Grapevine (Vitis spp.) is a woody, herbaceous tree-climbing plant. It was first grown in New Zealand in Kerikeri in the early 1800's, by a missionary, Sanuel Marsden (Danielmeier, 2008). A few years later, in 1832, James Busby who was the first Official British Resident of New Zealand produced the first wine in the same region. By the mid to late 1980s, New Zealand had started to gain recognition and reputation in the global wine industry, particularly for its 'Sauvignon Blanc'. The wine industry has grown to become a substantial participant of the New Zealand economy; the wine export value growing consistently to NZ$2.4 billion in 2023 (New Zealand Winegrowers, 2023). The most commonly planted varieties in New Zealand 'Sauvignon Blanc' (64.7% of the land), 'Pinot Noir' (13.6%), 'Chardonnay' (7.5%), and 'Pinot Gris' (6.7%) (New Zealand Winegrowers, 2023).

2.1.2 Winegrape Phenology

There are four accepted grapevine growth phenology systems. Baggiolini, (1952) proposed the first grapevine phenology system, which included 16 stages from bud break to leaf fall, in 1952. Eichhorn and Lorenz, (1977) then proposed a more comprehensive system consisting of 22 stages from "winter bud" to "end of leaf fall". This system is adopted by several wine industries. In addition, European Union developed the Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie (BBCH) system to standardize the phenological development stages of plants. In the adaptation of the BBCH system to grapevines, the stages are divided into seven macro-stages of growth, with each stage further broken down into numerous micro-stages. However, Coombe, (1995) found that the BBCH system contained terms that lacked universal understanding. Thus, Coombe developed the Modified Eichhorn and Lorenz system (Modified E-L) based on the grapevine phenology trials in Australia. The Modified E-L has 47 phases, which comprise a basic description of eight primary growth stages

and comprehensive intermediate stages. The physiology of main winegrape varietals in New Zealand can be described by:

Stage I. : Shoot and Inflorescence Development – occurring end of September to early November. This stage starts with the breaking of the dormant winter stage, marked by the sap rise. Four to six weeks following sap flow, the vines begin to generate foliage (tendrils and leaves) from the primary buds which formed at the end of previous year. Many leaves split off as the shoot grows to a height of about 10 cm, and the inflorescence is clearly visible.

Stage II. : Flowering – commencing around late November to early December. At this stage the young shoots start to separate into separate meristems and form flower clusters at the vine's nodes. Each flower within the cluster has a cap that, as it grows, falls off to reveal the inner floral components of the vine. It is considered to be in full bloom when 50% of the caps have fallen off.

Stage III. : Berry Development – from the end of flowering in mid-December to January. During this stage, the green berries grow at a right angle to the stem after setting on the clusters. Over the following two months, the berries grow in bulk, causing the clusters to hang down once they reach the size of peas. At this stage, minimal chemical variation is observed in the berry, aside from a subtle increase in acidity.

Stage IV. : Ripening – commencing around the end of January to early March. Following berry development, the berries go through a quick phase of chemical and physical change called veraison. At veraison, vegetative growth of shoots comes to an end, and the berries get softer, bigger, start to accumulate sugar and change color. Once the sugar/acid ratio and flavor profile reach their optimal levels, the berries are deemed to be harvest-ripe.

Stage V. : Senescence – from late March to late April. The vines enter their senescence stage after harvest. The canes become hard and woody, while the leaves fall from the vines.

Figure 2.1: Modified E-L system for grapevine phenological development stages (Coombe, 1995).

2.1.3 Winegrape Quality Attributes

2.1.3.1 Total Soluble Solids

Total soluble solids (TSS) is a measurement of soluble substances such as soluble sugars, pectin, organic acids, amino acids, and ascorbic acids, often expressed as a percentage (%) or in degrees Brix (°Brix) (Agulheiro‐Santos et al., 2022). In wine grapes, TSS are normally 99% sugars (glucose and fructose) (Wolf, 2008). Thus, TSS is a good assessment of grape sugar content. TSS can be easily measured using a manual or digital refractometer. The sugar content in wine grapes directly affects potential alcohol content after fermentation and the volume of residual sugars in the produced wine. As a result, during the grape harvest stage, the TSS value is crucial for determining the best harvest timing. In general, the recommended °Brix level of berry harvest-ripe is between 18-24 °Brix. This range varies depending on the grape varieties and the vine growing conditions (Wolf, 2008).

2.1.3.2 Acidity

In grape berries, acidity contributes to the tart or sour taste. Both insufficient and excessive levels of acidity are problems for winemaking. Inadequate acidity results in flat and uninteresting wines, while an excess of acidity results in wines that are overly sharp. Adequate acidity balances sweet and bitter flavors, contributing to a refreshing and crisp quality in the wine. In addition, optimal acidity also aids in stabilizing color and proteins, preventing the growth of spoilage microbes, and improving the efficacy of sulfur dioxide used in winemaking. While winemakers can add acid, receiving grapes with the desired acidity level simplifies the winemaking process and avoids additional expenses. Malic acid and tartaric acid are the dominant organic acids found in grape berries (Yinshan et al., 2017). . Titratable acidity (TA) and pH are two common methods to characterize acidity in grapes.

2.1.3.2.1 Titratable Acidity

Titratable acidity (TA) commonly used in the measurement of the organic acids present in grapes. Tartaric acid and malic acid are the principal organic acids found in grapes. TA

measures the total concentration of these organic acids and their salts, primarily representing the tartness and sourness in grapes and wine. It is important to note that TA is not a direct indicator of wine stability. In general, the recommended TA level of berry harvest-ripe is between 5-10 g/L tartaric acid equivalents This range vary depending on the grape varieties and the vine growing conditions.

2.1.3.2.2 pH

The pH is the concentration of associated or free hydrogen. This measure provides an indication of the intensity of the effective and active acidity in the wine, which is distinct from titratable acidity (TA) that measures the total concentration of all acids, including those that are bound or associated with other compounds. The pH level in grapes is important for chemical and microbial stability but has little direct impact on the taste profile of the grape and wine. As TA decreases, the pH generally tends to increase. The usual pH level for grape berries typically ranges between 2.9 and 4.0. This range can be influenced by grape varieties and the vine growing conditions.

2.2 Precision Viticulture

Precision viticulture (PV) is a modern technique, which use is growing rapidly in the ‘New and old word’ wine countries such as Australia, New Zealand, Canada and Europe (Baluja et al., 2013; Bramley, 2022). In PV, viticulturists detect and manage the spatial variation within vineyards, which enables them to assist growers in improving management and increasing grape yield and quality. (Baluja et al., 2013). Winemakers regard the uniformity of a bunch of grape berries as a key aspect of wine quality. This is because using uniform bunches can make wine with the same flavor and aroma. This can be more profitable than selling wines that have been downgraded because a relatively small part of the input grapes lack the desired characteristics (Bramley, 2022). Accordingly, the primary goal of modern vineyard management is to ensure that the grapes in the vineyard meet a specific grade.

Environmental variables vary spatially within vineyards, whereas in most vineyards the management strategy is uniform (van Leeuwen et al., 2008). As the result of uniform application, parts of the vineyard will receive less inputs than required, while other parts might receive more than they need. This results in grape size or quality being different in various harvest areas, which is problematic for winemakers. These factors have prompted researchers and growers to begin to adopt PV techniques. Due to advances in global positioning systems (GPS), geographical information system (GIS) and remote sensing (RS), viticulturists have more tools to measure and monitor systems and variability within vineyards. With these tools viticulturists can analyze the spatial and temporal variability of environmental variables, that affect grape yield and quality. Site-specific viticultural management practices can therefore be carried out.

Over the past decade, PV tools have been increasingly applied, and in many cases have been extensively studied for their aid to profitability and accuracy, with promising results (Bramley, 2022). Bramley et al., (2003) showed the economic advantages of targeted harvesting based on remotely sensed imagery in a 3.3 ha Australian vineyard were estimated to be AUD $2,661/t compared to normal harvesting. In addition, using vegetation index maps to delineate management zones is estimated to provide a financial saving of AUD $400/t in a commercial vineyard (Newson et al., 2012).

The development of spatial technologies plays a key role in the adoption of PV. It becomes possible and easy to collect information for mapping the canopy status, vine yield and soil maps. Remote sensing by satellite, unmanned aerial vehicle (UAV) and proximal sensors can provide reliable images about the land cover, canopy status and grape yield and quality (Kasimati et al., 2021; Wei et al., 2023, 2022). For example, Wei et al., (2022) used vegetation indices obtained from UAV to predict vine water status. In addition, an accurate yield map for grapes can be produced cheaply and conveniently by grape harvesters with differential global positioning systems instrumentation during the harvest (Taylor et al., 2005).

2.2.1 Spatial-temporal Variation in Vine Growth

Many researchers have studied the spatial and temporal variability of vine vegetative growth (vigor) in vineyards (Trought et al., 2008; Trought and Bramley, 2011; Bramley et al., 2011b, 2019). In these studies, they used the pruning mass, trunk circumference, trunk cross-sectional area and remote/proximal sensing images to assess vine vigor. Many studies showed that some vine vigor indicators such as: trunk circumference and pruning mass varied in the same timeframe within vineyard (Trought et al., 2008; Bramley et al., 2019). Bramley et al., (2011) measured trunk circumference of 2028 vines and used block kriging to interpolate the map. Although, the interpolation map showed very similar patterns of variation with the remote sensing images within vineyard, it took a lot of time to collect the 2028 samples in this study. Using remote sensing techniques to explore the spatial and temporal variation in vine vigor status can reduce the economic cost of sampling.

The two main vegetation indices normalized vegetation index (NDVI) and plant cell density (PCD), which are provided by remote or proximal sensing were used to assess the spatial-temporal variation of vine vigor within vineyards (Bramley, 2022). Compared with direct measurement, remote sensing images can provide high density sampling within a vineyard. This data is suitable for kriging to interpolate the map which can assess the spatial variation of vine vigor. The NDVI and PCD indices are an indicator of vine vigor, as they provide a measure of the photosynthetically active biomass (PAB) which can reflect different vine canopy statuses (Bramley, 2022). Therefore, they are often used as indicators of vine vigor (Bramley et al., 2011b, 2019). Bramley et al., (2011) inferred that the patterns of variation in vine vigor in a New Zealand vineyard were stable in time based on the PCD images which were obtained by remotely sensed digital multispectral imagery (~50 cm on-ground resolution). In the study of Bramley and Hamilton, (2007), there was a strong correlation between the spatial structure of PCD imagery and yield maps in a mechanical spur-pruned vineyard, whereas in a cane-pruned vineyard, the correlation between yield and vine vigor, measured using a proximal sensor was weak ($p<0.05$) (Bramley et al., 2011b). Remote sensing images have been shown to predict yield in mechanical spur pruned vineyards when the images are collected during veraision

(Bramley, 2022). Vegetation vigor measured either directly or remotely showed temporal and spatial variation within vineyards (Bramley et al., 2011b, 2019).

Besides, the spatial-temporal variation of other vine attributes which represent plant growth were confirmed to have the ability to influence changes to grape yield and quality within a vineyard (Arnó et al., 2012; Jasse et al., 2021). These vine attributes have the potential to reflect plant growth stress, such as the stem water potential (SWP) which is a reliable indicator of the water deficit in grapes during the growing season (Jasse et al., 2021). Jasse et al., (2021) used SWP as an indicator to assess the spatial-temporal variability of plant water status during the growing season. The result showed that the spatial and temporal variability of SWP was evident during the growing season, and these variations can have a negative impact on the grape yield. When compared with the plant water status, the spatial variability of grape fertilizer status is rarely mentioned, as the fertilizer requirement per hectare in the vineyard is much lower than most other plants or crops. However, a deficiency of nutrients can decrease vine vigor, or even cause plant death in extreme cases; and an excess of nutrient inputs can reduce the process of maturation and grape quality (Arnó et al., 2012). In most viticulture practices, farmers prefer to assess the grape nutrient status through analyzing soil properties and petioles within the vineyard, and the fertilizer management is uniform. But the variation of grape nutrients is not uniform within a vineyard. For example, the N content of vines showed significant differences in vine vigor areas based on the measurement of leaf petioles (King et al., 2014). Thus, it is important to explore the spatial variability of vine nutrient status to apply site-specific management.

2.2.2 Spatial-temporal Variation in Grape Quality

Grape quality is an important factor for viticulturists and winemakers to consider. Grape quality can determine the sensory attributes of wine and thus profit. Several studies have explored the spatial-temporal variation of grape quality within a vineyard, and their relationship with vine vigor and yield (Bramley, 2005; Reynolds et al., 2007; Tisseyre et al., 2008; Trought et al., 2008; Trought and Bramley, 2011). The study of Bramley, (2005) assessed multi-year

data in Australian vineyards and used geostatistics to analyze the spatial variability of grape quality. As a result, most of the grape composition showed a marked spatial structure within a vineyard and varied during seasons. However, when it turns to the spatial analysis, most indices exhibited a ‘moderate’ degree of spatial dependence. The variogram of some parameters such as pH and TSS in some years showed a large nugget effect (no spatial dependence of variables). Tisseyre et al., (2008) reported very similar results. They provided a plausible reason for this that the sampling strategies are designed with less determination of the short-range variation. Moreover, Tisseyre et al., (2008) inferred that the inter seasonal variation in quality indicators is mainly due to the different weather in different growing seasons. While the spatial pattern of grape quality parameters was stable over three consecutive seasons in a vineyard in Spain (Baluja et al., 2013). It is a novel outcome which differs from the previous works (Bramley et al., 2005; Tisseyre et al., 2008). Baluja et al., (2013) concluded that the similarities between the weather conditions in different seasons may have contributed to this inter-seasonal stability. An important issue in the above studies is the economic impact on viticulturists of destructive sampling in vineyards and the cost and time of laboratory analysis (Bramley, 2022). The number of manual sampling points must be reduced due to economic and time factors, while sparse data reduces the support for exploring spatial and temporal variability of grape quality (Webster and Oliver, 2007). Therefore, developing a non-destructive and rapid method can solve this issue.

2.2.3 Early Grape Quality Prediction

As Kasimati et al., (2021) highlighted, an ability to predict the final grape quality during the vine growing stages is highly valuable. The vine vigor parameters (trunk circumference, remote and proximal imagery) are commonly used to be an indicator to predict the grape yield and quality in many studies (Hall et al., 2011; Trought and Bramley, 2011; Bramley et al., 2019) Using remote or proximal sensor monitor vine canopy status is widely used in New Zealand, and it can save sampling time and reduce destructive sampling (Bramley et al., 2019). The vegetation index such as NDVI and PCD which calculated from remote sensing imagery is seen to have a strong relationship with the grape quality parameters (Hall et al., 2011; Trought and

Bramley, 2011; García-Fernández et al., 2021; Kasimati et al., 2021). Hall et al., (2011) showed the relationship between NDVI and yield and quality parameters are likely to have different levels at different phenological stages (from post-budbreak to maturity). Trought and Bramley, (2011) noted that the ability of using PCD to predict fruit composition is best early in the ripening of the grape. The timing of imaging is an important factor determining the direction and strength of correlations between vegetation index and grape quality. Kasimati et al., (2021) used the NDVI data which obtained from different platforms and growing stages to predict the grape quality parameters. The result showed that the NDVI data collected at berries pea-sized and veraison growth stages have strong relationship with the grape sugar content. In addition to NDVI and PCD, other vegetation spectral indices can also be used to reflect vine growing status and predict final quality. For example, García-Fernández et al., (2021), examined the correlation between 12 vegetation indices obtained from UAV imagery and grape quality parameters. The result showed a strong correlation between excess green index and grape sugar content.

2.3 VNIR-SWIR Spectroscopy in Wine Grape Quality Determination

Conventionally, grape growers start to assess the grape quality from veraison to harvest stage based on destructive tests, such as TSS by the refractometer, pH by pH meter and electrode. Grape growers select and measure around 200 berries per block and per variety to represent the vineyard's grape quality. However, previous studies have shown that grape quality varies spatially within vineyard or between cluster of grapes (Baluja et al., 2013; Bramley, 2005). Conventional measuring methods result in the destruction of the berries, while allowing only a small proportion of berry samples to be measured.

The use of VNIR-SWIR spectroscopy to determine grape quality in a non-destructive way can provide the possibility to assess quality on a large number of berries and determine multiple quality attributes on the same berries during harvest stages (Walsh et al., 2020). Spectroscopy involves examining how objects interact with electromagnetic radiation. When electromagnetic radiation interacts with an object, it carries information about the object's composition or the nature of processes occurring within it (Nicolai et al., 2007a). This information is generally contained in the electromagnetic radiation transmitted through the object or reflected from it and is revealed in the spectrally selective nature of the wavelengths absorbed or emitted by the object (Milton, 2009). Electromagnetic radiation can be classified as radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X-rays, and gamma waves based on the frequency and wavelength (Figure 2.2) (Lambert and Edwards, 2023). Whereas the actual instruments used to measure the interactions between electromagnetic and matter are referred to as "spectrometers" or "spectroradiometers" (Milton et al., 2009).

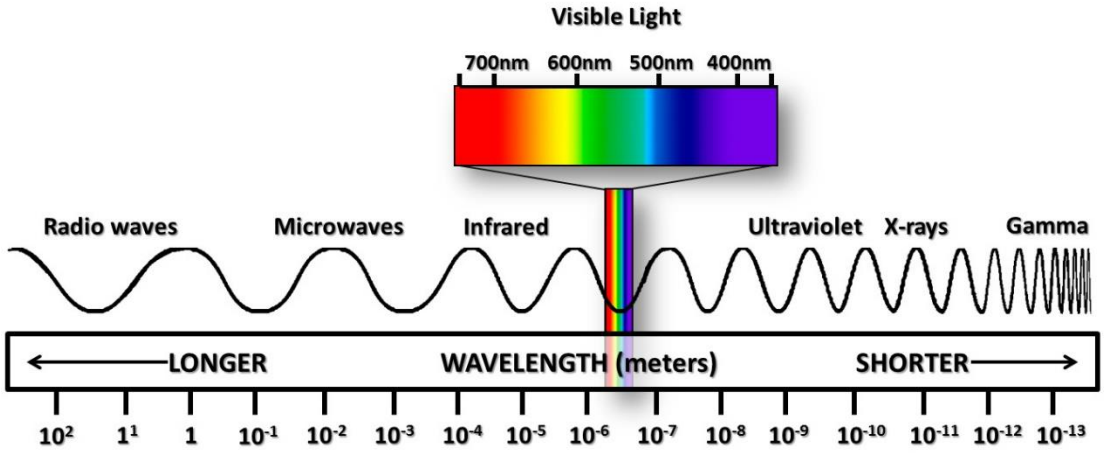

Figure 2.2: Electromagnetic spectrum (Lambert & Edwards, 2023)

VNIR-SWIR spectroscopy (ranging from 400 nm to 2500 nm) has been widely applied in assessing berry quality. Of this range, the 450-750 nm region is referred to as visible (Vis), 750-2500 nm region is considered the near infrared (NIR), the 1000-2500 nm region is also referred to as the shortwave infrared (SWIR) (Walsh et al., 2020). Visible spectroscopy involves

measure of the absorption of the light associated with electronic transitions of molecules. NIR spectroscopy is based on vibration of molecular bonds and occurs in the NIR range (Walsh et al., 2020).

2.3.1 Visible Spectroscopy

The pigment such as chlorophylls, carotenes, xanthophylls, flavonols, anthocyanins and other phenols determine the color of the fruit. Visible spectroscopy is widely used to determine the fruit pigments based on the pigments' spectral features. The absorbance peak in the wavelength ranging from 420 to 503 nm is related to carotenes and xanthophylls (Walsh et al., 2020). The strong absorbance peak near to 475 nm is related to beta-carotene. Absorbance in the blue region around 350-500 nm is caused by xanthophylls, lutein and violaxanthin (Walsh et al., 2020). In addition, the absorption feature at 365-430 nm is dependent on flavonols (Merzlyak et al., 2005). The anthocyanin pigment sugar-protein complex displays strong absorption across 530-550 nm (Merzlyak et al., 2005). Absorbance in the blue and red wavelengths region is caused by the chlorophylls (Kuai et al., 2018).

2.3.2 Near-Infrared Spectroscopy

Near-infrared (NIR) spectroscopy is widely used to measure the water, carbohydrates, and proteins level in fruit (Chandrasekaran et al., 2019). The chemical group such as C-H, O-H in these compounds are NIR-active chemical groups. When irradiated by NIR light, vibrational energy changes in the chemical group such as C-H, O-H, involving stretching or bending, lead to anharmonic vibrations, enabling overtone transitions and combination modes that manifest as specific absorption bands in the NIR region. The first, second and third overtones of O-H stretching contribute the absorption at 840, 960 and 1440 nm. Absorbance near 1900 nm related to a combination O-H features. The first, second and third overtones features of C-H contribute to absorption near 910, 1100, and 1700 nm. However, the absorption features are broad and overlapping, making the interpretation of the raw absorption spectra of fruit challenging (Walsh et al., 2020).

2.3.3 Acquisition Modes of VNIR-SWIR Spectroscopy

There are three common modes for VNIR-SWIR spectroscopy acquisition: reflectance, transmittance, and interactance modes (Figure 2.3). Most studies used reflectance mode for grape berry quality non-destructive estimation (Li et al., 2019). The position of light source, detector and fruit determines the different acquisition modes. In reflectance modes, the detector captured the specular and diffusely reflected light from the illuminated part of fruit (Walsh et al., 2020). The configuration of detectors and light sources in reflectance mode is typically set at a 45° angle to minimize specular reflections. This angle is chosen to reduce the impact of specular reflections, which do not provide information about the internal attributes of the fruit (Cozzolino et al., 2011). Reflectance mode is usually used to determine external quality features. In transmittance mode, the detector is placed in front of the light source and received the transmitted light through the fruit. Internal component concentration is typically detected using transmittance mode (Greensill and Walsh, 2000). However, transmittance mode has a low signal level from light attenuation and is influenced by the thickness of sample. In interactance mode, the light source and the detector are positioned on the same side of the sample and parallel to each other. In addition, a light barrier separates the light source and detector, preventing direct entry of light from specular reflection into the detector. Compared with reflectance mode, interactance mode can detect deeper information within the fruit and has less surface effects. Simultaneously, interactance mode minimizes the impact of sample thickness, making it a practical advantage over transmission mode (Walsh et al., 2020).

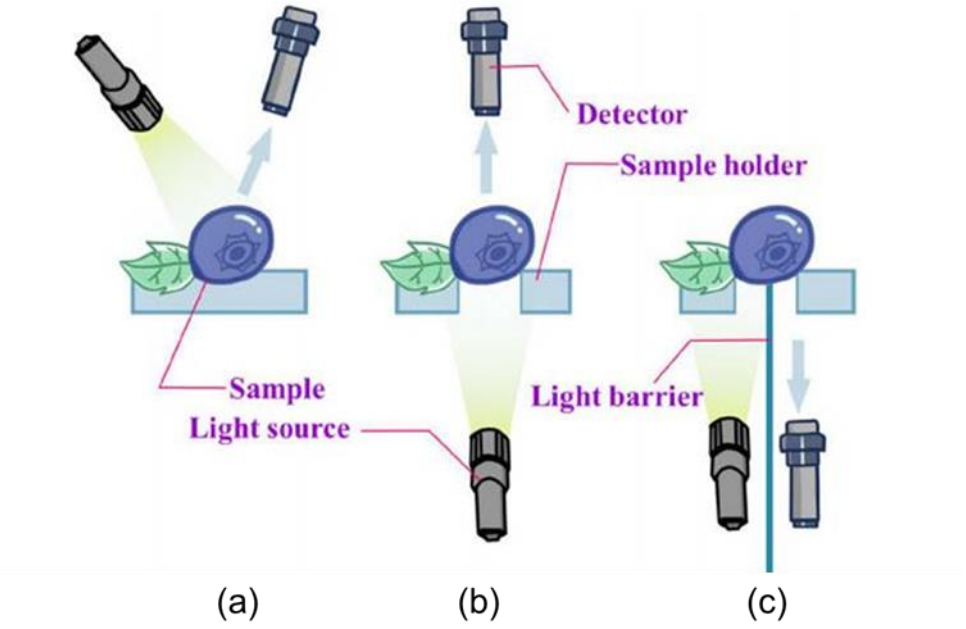

Figure 2.3: Acquisition modes of VNIR-SWIR spectroscopy. Reflectance mode (a), Transmittance mode (b), Interactance mode (c) (Li et al., 2019)

2.3.4 Applications of VNIR-SWIR Spectroscopy in Wine Grapes

For the wine grape industry, VNIR-SWIR spectroscopy is most commonly used in prediction of oenological parameters by developing regression models (Table 2.1). For instance, using VNIR-SWIR spectroscopy under lab conditions can achieve rapid and non-destructive prediction of wine grape quality attributes with high prediction accuracy of R^2 above 0.9 (dos Santos Costa et al., 2019; Urraca et al., 2016). However, González-Caballero et al., (2012) used VNIR-SWIR spectroscopy (380-1700 nm) to predict the TSS, pH, and reduced sugar content of 28 different wine grape varieties under lab conditions and obtained a moderate prediction accuracy (Table 2.1) Compared with dos Santos Costa et al., (2019) and Urraca et al., (2016), González-Caballero et al., (2012) applied spectroscopy to 108 grape bunches, where the small sample size probably affected the prediction accuracy.

Wine grapes are non-climacteric fruits, exhibit no further ripening once removed from the vine. While in transit to the lab for spectral measurements, grape berries tend to lose their fluids,

dehydrate, and become sensitive to microbial decay, which affects subsequent spectral and quality measurements. Some studies attempt to predict wine grape quality attributes based on on-site VNIR-SWIR spectral data in situ (Larraín et al., 2008; González-Caballero et al., 2012; Urraca et al., 2016; Kalopesa et al., 2023). For the studies conducted using VNIR-SWIR spectroscopy directly on-site, the prediction accuracy was lower than those under lab conditions (Table 2.1). Compared with controlled lab conditions, it is necessary to consider the influence of sunlight and measurement angle on spectral measurement when applying VNIR-SWIR spectroscopy on-site. Larraín et al., (2008) used a spectrometer (640-1300 nm) with a controlled light source to non-destructive measure the wine grape berry TSS, pH and anthocyanin content when berries were still on the hanging clusters of the vines. The instrument used in their study possesses its own light source, and this light source is maintained at a specific angle to the spectrometer. This setup ensures a stable measurement condition, consequently leading to a good prediction performance (Table 2.1). Kalopesa et al., (2023) used a similar instrument, but the spectrum ranged from 350 to 2500 nm to non-destructive measure grape TSS values in-field. The result demonstrated that the VNIR-SWIR spectroscopy have the potential to predict the TSS value of Malagouzia, Sauvignon Blanc, and Syrah varieties ($R^2 > 0.84$). However, when using VNIR-SWIR spectroscopy to predict the TSS value of Chardonnay, a moderate prediction performance was achieved with an R^2 value of 0.63.

Table 2.1: Applications of VNIR-SWIR spectroscopy in wine grapes based on reflectance mode

Reference	Spectra Range (nm)	Cultivar	Applied to	Oenological Parameters	Sample Number	On-site	Method	Result
(Kalopesa et al., 2023)	350-2500 nm	Chardonnay	Berry	TSS	139	√	Regression	R^2 = 0.63, RMSE = 2.1 °Brix
(Kalopesa et al., 2023)	350-2500 nm	Malagouzia	Berry	TSS	179	√	Regression	R^2 = 0.84, RMSE = 1.96 °Brix
(Kalopesa et al., 2023)	350-2500 nm	Sauvignon Blanc	Berry	TSS	195	√	Regression	R^2 = 0.86, RMSE = 2.20 °Brix
(Kalopesa et al., 2023)	350-2500 nm	Syrah	Berry	TSS	230	√	Regression	R^2 = 0.87, RMSE = 1.76 °Brix
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Grape Skin	Anthocyanin	84	-	Regression	RMSE = 0.81 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Grape Skin	Flavanols	84	-	Regression	RMSE = 0.68 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Grape Skin	Flavonols	84	-	Regression	RMSE = 0.15 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Grape Skin	Phenolic Acids	84	-	Regression	RMSE = 0.11 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Grape Skin	Total Phenolic Compounds	84	-	Regression	RMSE = 1.30 mg/g

(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Berry	Anthocyanin	84	-	Regression	RMSE = 1.09 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Berry	Flavanols	84	-	Regression	RMSE = 0.72 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Berry	Flavonols	84	-	Regression	RMSE = 0.12 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Berry	Phenolic Acids	84	-	Regression	RMSE = 0.13 mg/g
(Ferrer-Gallego et al., 2011)	1100-2500 nm	Graciano	Berry	Total Phenolic Compounds	84	-	Regression	RMSE = 1.26 mg/g
(dos Santos Costa et al., 2019)	350-2500 nm	Syrah	Berry	TSS	432	-	Regression	R^2 = 0.97, RMSE = 0.89 °Brix
(dos Santos Costa et al., 2019)	350-2500 nm	Syrah	Berry	Anthocyanin	432	-	Regression	R^2 = 0.94, RMSE =17.54 cg/kg
(dos Santos Costa et al., 2019)	350-2500 nm	Syrah	Berry	Yellow Flavonoids	432	-	Regression	R^2 = 0.76, RMSE =8.54 cg/kg
(dos Santos Costa et al., 2019)	350-2500 nm	Cabernet Sauvignon	Berry	TSS	576	-	Regression	R^2 = 0.97, RMSE = 1.08 °Brix
(dos Santos Costa et al., 2019)	350-2500 nm	Cabernet Sauvignon	Berry	Anthocyanin	576	-	Regression	R^2 = 0.97, RMSE =15.11 cg/kg
(dos Santos Costa et al., 2019)	350-2500 nm	Cabernet Sauvignon	Berry	Yellow Flavonoids	576	-	Regression	R^2 = 0.73, RMSE =13.88 cg/kg

(dos Santos Costa et al., 2019)	350-2500 nm	Syrah & Cabernet Sauvignon	Berry	Maturation Stages	1008	-	Classification	Accuracy = 93.15%, False Positive Ratio = 4.83
(Urraca et al., 2016)	1600-2400 nm	Tempranillo	Berry	TSS	1600	-	Regression	R^2 = 0.91, RMSE = 1.42 °Brix
(Urraca et al., 2016)	1600-2400 nm	Tempranillo	Berry	TSS	43	√	Regression	R^2 = 0.38, RMSE = 1.68 °Brix
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	TSS	71	√	Regression	R^2 = 0.82, RMSE = 1.48 °Brix
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	TA	71	√	Regression	R^2 = 0.81, RMSE = 1.48 g tart. acid dm^{-3}
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	pH	71	√	Regression	R^2 = 0.81, RMSE = 0.15
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	Anthocyanin	71	√	Regression	R^2 = 0.71, RMSE = 133.90 mg dm^{-3}
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	Polyphenols content	71	√	Regression	R^2 = 0.68, RMSE = 4.74
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	TSS	71	√	Classification	Accuracy = 77.1%, False Negative = 8
(Guidetti et al., 2010)	400-1000 nm	Nebbiolo	Berry	TA	71	√	Classification	Accuracy = 68.6%, False Negative = 11

(González-Caballero et al., 2010)	380-1700 nm	28 varieties	Bunch	TSS	108	-	Regression	R^2 = 0.57
(González-Caballero et al., 2010)	380-1700 nm	28 varieties	Bunch	pH	108	-	Regression	R^2 = 0.51
(González-Caballero et al., 2010)	380-1700 nm	28 varieties	Bunch	Reducing-sugar content	108	-	Regression	R^2 = 0.61
(González-Caballero et al., 2012)	1600-2400 nm	Pedro Ximénez Cabernet Sauvignon	Bunch	Ripening stage	24	√	Classification	Accuracy = 79%,
(Larraín et al., 2008)	640-1300 nm	Cabernet Sauvignon	Berry	TSS	724	√	Regression	R^2 = 0.93, RMSE = 1.01 °Brix
(Larraín et al., 2008)	640-1300 nm	Cabernet Sauvignon	Berry	pH	678	√	Regression	R^2 = 0.761, RMSE = 0.130
(Larraín et al., 2008)	640-1300 nm	Cabernet Sauvignon	Berry	Anthocyanin	205	√	Regression	R^2 = 0.618, RMSE = 0.249
(Larraín et al., 2008)	640-1300 nm	Carménère	Berry	TSS	740	√	Regression	R^2 = 0.923, RMSE = 1.218 °Brix
(Larraín et al., 2008)	640-1300 nm	Carménère	Berry	pH	685	√	Regression	R^2 = 0.798, RMSE = 0.137

(Larraín et al., 2008)	640-1300 nm	Carménère	Berry	Anthocyanin	218	√	Regression	R^2 = 0.637, RMSE = 0.315
(Larraín et al., 2008)	640-1300 nm	Merlot	Berry	TSS	144	√	Regression	R^2 = 0.921, RMSE = 1.099 °Brix
(Larraín et al., 2008)	640-1300 nm	Merlot	Berry	pH	135	√	Regression	R^2 = 0.771, RMSE = 0.141
(Larraín et al., 2008)	640-1300 nm	Merlot	Berry	Anthocyanin	135	√	Regression	R^2 = 0.681, RMSE = 0.300
(Larraín et al., 2008)	640-1300 nm	Pinot Noir	Berry	TSS	145	√	Regression	R^2 = 0.875, RMSE = 1.267 °Brix
(Larraín et al., 2008)	640-1300 nm	Pinot Noir	Berry	pH	135	√	Regression	R^2 = 0.56, RMSE = 0.159
(Larraín et al., 2008)	640-1300 nm	Pinot Noir	Berry	Anthocyanin	135	√	Regression	R^2 = 0.398, RMSE = 0.183
(Larraín et al., 2008)	640-1300 nm	Chardonnay	Berry	TSS	409	√	Regression	R^2 = 0.874, RMSE = 1.089 °Brix
(Larraín et al., 2008)	640-1300 nm	Chardonnay	Berry	pH	389	√	Regression	R^2 = 0.661, RMSE = 0.088

2.4 Hyperspectral Imaging System in Wine Grape Quality Determination

Spectroscopy techniques can provide high spectral resolution data, however, which does not give information on spatial resolution of traits in grape berries (Manley, 2014). This limitation restricts their application in quantifying spatial distribution and structure-related attributes. Hyperspectral imaging systems, can provide both spectral and spatial information in one system for the quality determination of wine grapes (Pu et al., 2015). Hyperspectral imaging systems capture images containing numerous uninterrupted wavelengths. Every pixel in the hyperspectral image, which is composed of dozens or hundreds of images, has a unique spectrum spanning a continuous wavelength range. (Wu and Sun, 2013).

2.4.1 Principles of Hyperspectral Imaging System

A hyperspectral image is a three-dimensional (3D) dataset that combines spectral information (wavelength λ) with two-dimensional spatial information (x rows and y columns) (Wu and Sun, 2013). As an example, the hyperspectral cube of grape berries acquired using Specim hyperspctral imagaing system is illustrated in Figure 2.4. The hyperspectral cube is comprised of consecutive sub-images of the same object, where each one of them is arranged along different wavelengths. The hyperspectral image I (x, y, λ) can be be interpreted either as distinct spatial images I(x, y) at each individual wavelength (λ), or as a spectrum I(λ) at each individual pixel (x, y). As hyperspectral imaging captures spatially distributed spectral responses at the pixel level, it enables the flexible selection of any regions of interest on a target object.

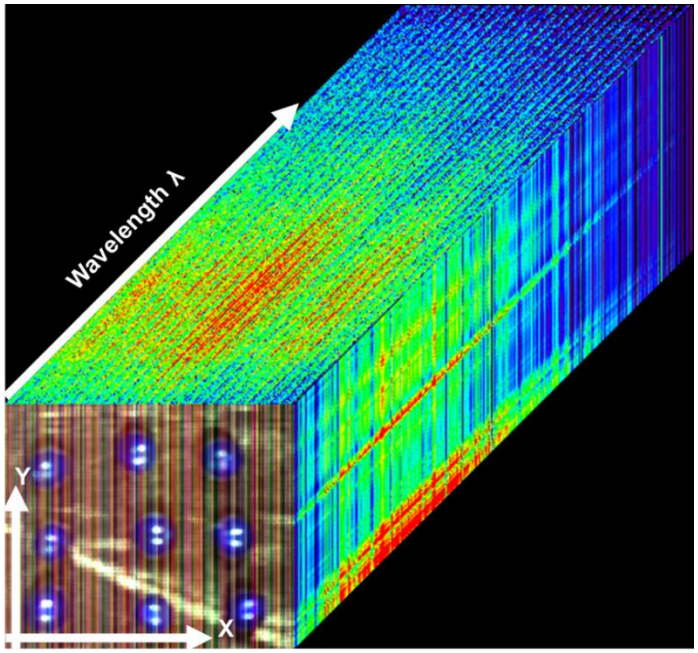

Figure 2.4: Hyperspectral cube of wine grapes

2.4.2 Acquisition of Hyperspectral Imaging

There are four ways to capture hyperspectral imaging: whiskbroom (point scanning), push broom (line scanning), spectral scanning (area scanning or plane scanning), and snapshot (single shot) (Wu and Sun, 2013). In the whiskbroom approach, hyperspectral cameras scan one single point and provide the spectrum of this point. When the camera or the sample moves along two spatial dimensions (x and y), other points are scanned. This approach provides the hyperspectral cube in the band-interleaved-by-pixel (BIP) format (Sun, 2010). The advantages of whiskbroom are very high spectral resolution. While this approach is very time-consuming during the image acquisition (Elmasry et al., 2012; Wu and Sun, 2013). In the push broom approach, hyperspectral cameras capture a whole line of an image and spectral information simultaneously corresponding to each spatial pixel in line (Wu and Sun, 2013). The full image can be obtained as the line is scanned continuously across the sample in one direction. This approach provides the hypercube in the band-interleaved-by-line (BIL) format (Gowen et al., 2007). The push broom method is commonly designed for a conveyor belt system and used in

food process lines. Although the push broom method is faster than whiskbroom, the exposure time can be set at only one value for all wavelengths, which results in spectrum saturation during spectral measurement. The whiskbroom and push broom are spatial scanning methods. Compared with the spatial scanning method, the spectral scanning method captures the entire 2-D monochrome image (x, y) for a given wavelength at one time (Wu and Sun, 2013). By repeating this scan across the entire spectral range, a complete hypercube can be obtained. This approach provides the hypercube in the band sequential (BSQ) format. The spectral scanning method can set the exposure time based on different wavelengths. Additionally, this method does not need either the sample or hyperspectral camera to move and is suitable for applying to stationary samples. This method is not suitable for moving samples and real-time delivery inspection. Finally, the snapshot method captures the spatial and spectral information of the full image simultaneously. This method is very fast, although it only can provide limited spectral and spatial resolutions (Wu and Sun, 2013).

2.4.3 Applications of Hyperspectral Imaging System in Wine Grapes

Figure 2.5: Acquisition mode of hyperspectral imaging. Point scanning (a), Line scanning (b), Area scanning (c), Single shot (d) (Wu & Sun, 2013)

Numerous researchers have explored the possibility of hyperspectral imaging system in assessing various grape quality attributes, including: TSS, pH, titratable acidity, and

anthocyanins content (Table 2.2). The majority of these studies were performed in laboratory settings with controlled conditions. For example, Gomes et al., (2021b) utilized line-scanning hyperspectral imaging system in the Visible and NIR region (400-1000 nm) to predict pH, TSS and anthocyanin contents in wine grape berries under controlled laboratory conditions. Gomes et al., (2021b) showed that hyperspectral imaging system have the potential to predict different grape quality attributes, with RMSE values of 0.168 for pH, RMSE values of 0.998 °Brix for TSS, and RMSE values of 19.733 mg/l for anthocyanin contents based on different regression models. Regarding studies conducted directly in the field, Benelli et al., (2021) and Gutiérrez et al., (2019) used an on-the-go hyperspectral imaging system to predict grape maturity. Gutiérrez et al., (2019) achieved high prediction accuracy, with R^2 values of 0.92 (RMSE = 1.274 °Brix) for TSS and R^2 values of 0.83 (RMSE = 0.211 mg g^{-1}) for anthocyanin based on support vector regression model (SVR). Benelli et al., (2021) also achieved good prediction accuracy, with R^2 values of 0.77 (RMSE = 0.79 °Brix) for TSS using a partial least squares regression (PLSR) model. In addition, Tsakiridis et al., (2023), used a snapshot hyperspectral camera to predict gape sugar content, and resulted in high prediction performance with R^2 values of 0.70.

The equipment used in most studies operates from 400 to 1000 nm, as the equipment capable of operating at wavelength above 1000 nm are more expensive. However, the overtone features in the region between 1100-2500 nm are related to fruit quality attributes. Chen et al., (2015) used hyperspectral imaging system in the NIR region (900-1700 nm) to predict anthocyanin content under the lab condition, with R^2 of 0.94. In addition, other study employed the spectrometer in 1000-2500 nm to predict grape quality attributes (Table 2.1). Compared with spectrometer, hyperspectral imaging systems can provide spatial resolution of grape berries. This capability is particularly valuable when analyzing objects with spatial variations, such as grape clusters.

1 Table 2.2: Applications of hyperspectral imaging system in wine grapes based on reflectance mode

Reference	Spectra Range (nm)	Cultivar	Applied to	Oenological Parameters	Sample Number	On-site	Method	Result
(Tsakiridis et al., 2023)	450-1000 nm	Chardonnay	Berry	TSS	57	√	Regression	R^2 = 0.74, RMSE = 1.66 °Brix
(Tsakiridis et al., 2023)	450-1000 nm	Malagouzia	Berry	TSS	71	√	Regression	R^2 = 0.62, RMSE = 1.72 °Brix
(Tsakiridis et al., 2023)	450-1000 nm	Sauvignon Blanc	Berry	TSS	53	√	Regression	R^2 = 0.74, RMSE = 2.29 °Brix
(Tsakiridis et al., 2023)	450-1000 nm	Syrah	Berry	TSS	55	√	Regression	R^2 = 0.69, RMSE = 1.66 °Brix
(Silva and Melo-Pinto, 2021)	380-1028 nm	Touriga Franca	Berry	TSS	240	-	Regression	R^2 = 0.9690, RMSE = 0.92 °Brix
(Gomes et al., 2021a)	380-1028 nm	Touriga Franca Touriga Nacional Tinta Barroca	Berry	TSS	2665	-	Regression	R^2 = 0.948, RMSE = 1.18 °Brix
(Gomes et al., 2021a)	380-1028 nm	Touriga Franca Touriga Nacional Tinta Barroca	Berry	pH	2665	-	Regression	R^2 = 0.728, RMSE = 0.183

(Fernandes et al., 2015)	380-1028 nm	Touriga Franca	Berry	TSS	240	-	Regression	R^2 = 0.92, RMSE = 0.95 °Brix
(Fernandes et al., 2015)	380-1028 nm	Touriga Franca	Berry	pH	240	-	Regression	R^2 = 0.73, RMSE = 0.18
(Fernandes et al., 2015)	380-1028 nm	Touriga Franca	Berry	Anthocyanin	240	-	Regression	R^2 = 0.95, RMSE = 14 mg/l
(Gomes et al., 2017a)	380-1028 nm	Touriga Franca	Berry	TSS	324	-	Regression	R^2 = 0.95, RMSE = 1.34 °Brix
(Gomes et al., 2017b)	380-1028 nm	Touriga Franca Touriga Nacional Tinta Barroca	Berry	pH	465	-	Regression	R^2 = 0.723, RMSE = 0.170
(Gomes et al., 2017b)	380-1028 nm	Touriga Franca Touriga Nacional Tinta Barroca	Berry	Anthocyanin	465	-	Regression	R^2 = 0.906, RMSE = 22.1 mg/l
(Gomes et al., 2021b)	380-1028 nm	Touriga Franca	Berry	pH	240	-	Regression	RMSE = 0.168
(Gomes et al., 2021b)	380-1028 nm	Touriga Franca	Berry	TSS	240	-	Regression	RMSE = 0.998 °Brix
(Gomes et al., 2021b)	380-1028 nm	Touriga Franca	Berry	Anthocyanin	240	-	Regression	RMSE = 19.773 mg/l

(Benelli et al., 2021)	400-1000 nm	Sangiovese	Bunch	TSS	143	√	Regression	R^2 = 0.768, RMSE = 0.79 °Brix
(Benelli et al., 2021)	400-1000 nm	Sangiovese	Bunch	TSS	143	√	Classification	Area Under the Curve (AUC) = 0.96
(Lafontaine et al., 2015)	400-2500 nm	12 red varieties	Berry	TSS	3228	-	Regression	RMSE = 0.93 °Brix
(Gutiérrez et al., 2019)	400-1000 nm	Tempranillo	Bunch	TSS	144	√	Regression	R^2 = 0.92, RMSE = 1.274 °Brix
(Gutiérrez et al., 2019)	400-1000 nm	Tempranillo	Bunch	Anthocyanin	144	√	Regression	R^2 = 0.72, RMSE = 0.282 mg/g
(Chen et al., 2015)	900-1700 nm	Cabernet Sauvignon	Berry	Anthocyanin	120	-	Regression	R^2 = 0.94, RMSE = 0.0046 mg/g
(Silva et al., 2018)	380-1028	Touriga Franca Touriga Nacional Tinta Barroca	Berry	TSS	1388	-	Regression	R^2 = 0.90, RMSE = 3.19 °Brix
(Silva et al., 2018)	380-1028	Touriga Franca Touriga Nacional Tinta Barroca	Berry	pH	1388	-	Regression	R^2 = 0.81, RMSE = 0.25
(Silva et al., 2018)	380-1028	Touriga Franca Touriga Nacional	Berry	Anthocyanin	1388	-	Regression	R^2 = 0.89, RMSE = 35.6 mg/L

		Tinta Barroca						

2

2.5 Multivariate Statistical Analysis Applied to VNIR-SWIR Spectroscopy and Hyperspectral Imaging System

Multivariate statistical analysis of the spectral data extracted from spectrometers or hyperspectral imaging system should be able to achieve rapid and non-destructive predictions of the chemical or physical properties of the fruit (Cozzolino et al., 2011). Effective multivariate statistical analysis techniques can improve the predictive ability. Multivariate statistical analysis techniques are required to pre-process spectral data, extract important spectral information, and develop models (Nicolai et al., 2007a).

2.5.1 Spectral Pre-processing Method

Spectral pre-processing is the most important step in multivariate statistical analysis. The recorded sample spectral can be significantly affected by the light scattering, sample geometry and variation in temperature etc (Li, 2017). These unwanted variations during spectral measurement can be observed as baseline shifts and non-linearities (Rinnan et al., 2009). Effectively spectral pre-processing techniques can remove these unwanted variations.

2.5.1.1 Transformation

Spectroscopic measurements performed in reflectance or transmittance mode are usually converted to the absorbance according to the empirical Lambert-Beer's Law (Equation 2.1 – 2.2) (Rinnan et al., 2009). Lambert-Beer's law suggests that there is a linear relationship between the absorbance of the spectra and concentration of absorbing analyte (Nicolai et al., 2007a). Lambert-Beer's law is strictly valid only for dilute and non-scattering solution. However, in the case of fresh fruit, the presence of scattering complicates the application of Lambert-Beer's law. When light interacts with biological tissues in fresh fruits, it not only gets absorbed but also scattered. Scattering increases the effective pathlength travelled by light through the sample, leading to deviations from the predictions of the Lambert-Beer's law. Overall, the method of transforming reflectance or transmittance to absorbance does not seem to offer a large advantage compared to the untransformed spectra (Nicolai et al., 2007b).

$$A_\lambda = -\log_{10}(T) = \varepsilon_\lambda lc \quad \text{Equation 2-1}$$

$$A_\lambda = -\log_{10}(R) \cong \varepsilon_\lambda lc \quad \text{Equation 2-2}$$

where A is the absorbance, ε is the molar absorptivity, l is the pathlength of the light through the sample, c is the concentration of absorbing analyte, λ is the specific wavelength.

2.5.1.2 Scatter Correction Methods

Scatter correction methods are used to reduce the physical variability of spectral data and correct the baseline shits. The common scatter correction methods include Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) and normalization. MSC is the most common spectra pre-processing method and was developed by Martens et al., (1983). MSC attempts to remove the undesirable scatter effect by linearizing each spectrum to the reference spectrum of the sample. In most practices, the average spectrum of the sample is used as reference spectrum. SNV involves subtracting the mean spectrum from each individual spectrum and then dividing by the standard deviation of the spectrum. This process standardizes the shape of the spectra, making them more comparable and removing baseline shift. Normalization was used to standardize the intensity values across spectra. The primary aim is to reduce unwanted sources of variability related to differences in overall intensity levels, allowing for a more meaningful comparison and analysis of spectral features.

2.5.1.3 Spectral Derivatives

Spectral derivatives pre-processing methods are used to remove both additive and multiplicative effects in the spectra. The common spectral derivatives include first derivative and second derivative. The former are used to remove the baseline trend, while the latter are used to remove both baseline and linear trend. In addition, the spectral derivatives can combine with smoothing techniques in order to reduce the signal-to-noise ratio in the spectra. The most common smoothing methods is Savitzky-Golay filter. The Savitzky-Golay filter involves fitting a polynomial to the spectrum within a wavelength interval using the least-squares method

(Williams and Norris, 1987). However, smoothing techniques may remove valuable information which may not be clear at the time of removal (Li, 2017).

2.5.2 Variable Selection

Variable selection, also known as variable reduction, is developed to find a subset of variables from the entire range of spectra for simpler and robust model development (Xiaobo et al., 2010). Using informative variables can improve model performance, interpret models easier, and save computational cost (Cai et al., 2008). In addition, reducing the number of variables can be especially advantages in datasets that have a high number of variables in comparison to the number of samples available. Principal component analysis (PCA) is the most popular approach to reduce the number of variables (Silva et al., 2018; Silva and Melo-Pinto, 2021). Several other approaches have also been implemented in variable selection including correlation coefficient filtering, genetic algorithms, recursive feature elimination and interval partial least squares (Cai et al., 2008; Wei et al., 2021).

Another simple method of reducing the number of variables is by down sampling the original spectral data at a low spectral resolution. The spectrometers and hyperspectral imaging systems used in research area generally have a spectral resolution less than 5 nm. However, in most industry application, 10 nm spectral resolution is often sufficient (Nicolai et al., 2007a). High spectral resolution may have redundant information and increase the computational cost. Nicolai et al., (2007b) explored the prediction performance of NIR regression model based on different spectral resolution and found that a 5 nm spectral resolution can provide the best performance.

2.5.3 Regression and Classification Algorithms

Regression algorithms are used to establish a relationship between response variable (quality attributes of interest, such as TSS) and the explanatory variable (spectral value) (Maulud and Abdulazeez, 2020). The regression algorithms can be divided into linear regression and non-linear regression algorithms. The linear regression algorithms include multivariate linear regression (MLR), principal component regression (PCR) and partial least

squares regression (PLSR). PLSR is probably the most widely used linear regression algorithm for non-destructive fruit prediction studies. PLSR was introduced by Herman Wold to overcome multicollinearity in the 1960's (Wold et al., 2001). In addition, PLSR can effectively address dimensionality when the number of explanatory variables is greater than the number of samples. PLSR constructs the latent variables as a function of both explanatory and response variables. The required number of latent variables is optimized by the cross-validation procedure (Wold et al., 2001).

Classification algorithms aim to assign data to predefined categories or classes based on their feature characteristics (Moreda et al., 2009). The classification algorithms can be divided into supervised or unsupervised (Kim et al., 2000). Supervised classification algorithms are widely used for fruit quality discrimination. In the supervised classification algorithms, the label of each sample is known. The traditional classification algorithms include linear discriminant analysis (LDA) and partial least-squares discriminant analysis (PLS-DA).

Recent development of computational power and resources have made it easier to implement regression and classification tasks based on machine learning and deep learning algorithms. Machine learning and deep learning algorithms are promising techniques to process spectroscopic data or high spatial resolution hyperspectral imaging data. Generally, machine learning algorithms used in fruit quality detection studies include regularized regression, K-nearest neighbors (KNN), support vector machines (SVM) and tree-based models.

Most machine learning algorithms are referred to as shallow models, as they use one or two layers of data transformation to find the optimal representation (Janiesch et al., 2021). However, it is not always possible to find the optimal output representation based on a shallow model. Deep learning provides a multi-layer approach to capture complex patterns and hierarchical structures in high-dimensional data (Zhang et al., 2023). With the development of computational resources, deep learning has been widely applied in various fields, such as image analysis, natural language processing, spectral analysis, and speech recognition (Zhang

et al., 2021). There are many types of deep learning algorithms available, such as convolutional neural networks (CNNs), recurrent neural network (RNN), and long short-term memory neural network (LSTM). Multilayer perceptron (MLP) is the fundamental of all these methods (Zhang et al., 2023).

The layers of interconnected nodes are the building blocks of the MLP architecture. MLP is composed of three layers: input layer, hidden layer, and output layer. The input layer receives the original input features (Noriega, 2005). One or more hidden layers are placed between the input and output layer. and process the input information, which is the majority of the network. The output layer is the final layer in the network and produces the regression or classification result based on the processed information from the hidden layer (Figure 2.6).

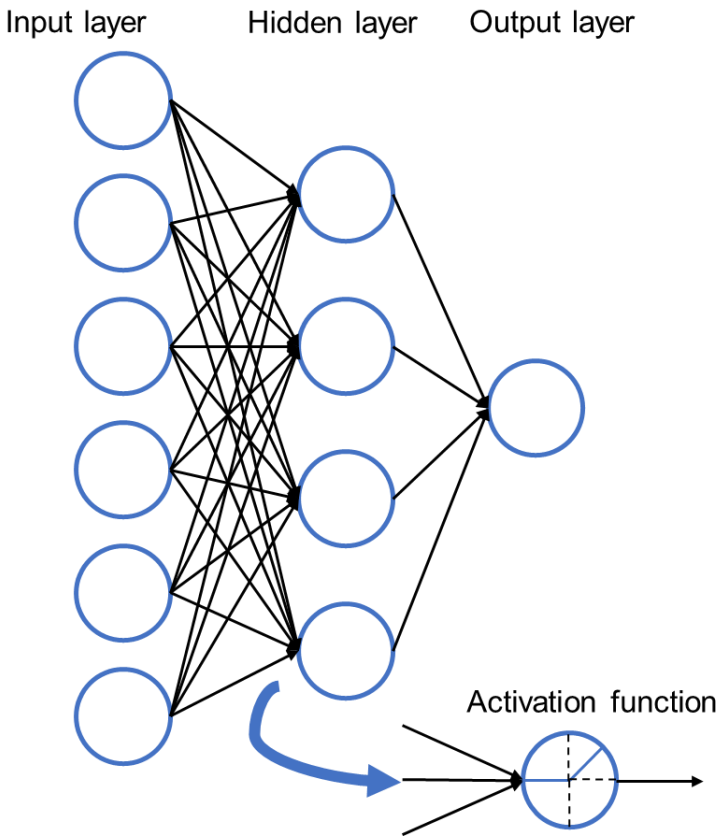

Figure 2.6: A simple multilayer perceptron architecture

The layers and nodes can decide the complexity of the MLP. More layers and nodes in the MLP can improve the model's capacity to learn new features. There is no one best method for selecting the number of hidden layers and nodes. Thus, the ideal number of hidden layers and nodes are the hyperparameters that need to be tuned when building the model. In most cases, 2-5 hidden layers are often sufficient for tabular data (Zhang et al., 2023). The rule of thumb is that choose more layers rather than fewer for optimal results. The number of nodes used in these hidden layers is determined by the number of features in the data (equal to or less than the number of features). It should be noted that the number of hidden layers and nodes in the MLP will affect the computational time. The activation function is another key hyperparameter in the neural network architecture other than the node and layer. The activation function serves as a mathematical function that decides if there is a sufficient amount of informative input at a node to trigger the transmission of a signal to the next layer (Figure 2.6). The common activation functions include:

$$\text{Rectified linear unit (ReLU)}: f(x) = \begin{cases} 0, for\ (x\ < 0) \\ x, for\ (x \geq 0) \end{cases} \quad \text{Equation 2-3}$$

$$\text{Sigmoid}: f(x) = \frac{1}{1+e^{-x}} \quad \text{Equation 2-4}$$

$$Tanh: f(x) = \frac{1-e^{2x}}{1+e^{-2x}} \quad \text{Equation 2-5}$$

MLP can provide a good performance when processing regular tabular data (the data consists of rows corresponding to examples and columns corresponding to features). However, when dealing with image data, MLP does not properly consider the spatial structure between features (Zhang et al., 2023). In addition, MLP needs to calculate billions of parameters when processing image data, which requires very high computing performance. Thus, convolutional neural networks (CNNs) were developed for image processing (Naranjo-Torres et al., 2020). Compared to MLP, CNNs reduce the number of parameters through two key theories: locality and weight sharing (LeCun et al., 1995). Locality means that in CNNs, each neuron in a convolutional layer only needs connect to a local region of the input data. This allows each neuron to focus on a small part of the input, rather than the entire input. This way, the network

captures local features instead of entire features, reducing the overall number of parameters. In CNNs, the same set of weights is used to process different parts of the input. Weight sharing effectively reduces the number of parameters that need to be learned, as the weights are not learned separately for each position but are shared. The fundamental building block in CNNs is the convolutional layer. In the convolutional layer, the kernel, padding, and stride determine the convolution operation (Figure 2.7). The kernel is a matrix which slides over the input image to extract the feature. The size of the kernel determines the specific region of the input image that the kernel considers at a time. Padding is used to add extra pixels of filler around the boundary of the input image before the convolution operation. Padding can prevent the loss of the edge information of the input image. The padding value is often zero. The stride is the step size with which the kernel moves across the input data (Figure 2.7). It determines how much the kernel shifts at each step. A larger stride reduces the spatial dimensions of the output feature map, as the kernel skips more pixels at each step. A smaller stride retains more spatial information but may increase computational cost. When we process multispectral or hyperspectral image which have multiple channels, it is necessary to construct the same number of kernel windows as the image channels. The pooling layer is another important block in CNNs. A pooling layer is used to down-sample the spatial dimensions of the input image. Pooling can save the computational time of the CNNs and control overfitting, while retaining important features. There are two main types of pooling layers: max pooling and average pooling. Pooling layers have a pooling window that is slid over all regions in the input image. Compared with convolutional layers, pooling operates only calculate the maximum or the average value of the elements in the pooling window. As with convolutional layers, pooling layers also contain padding and stride.

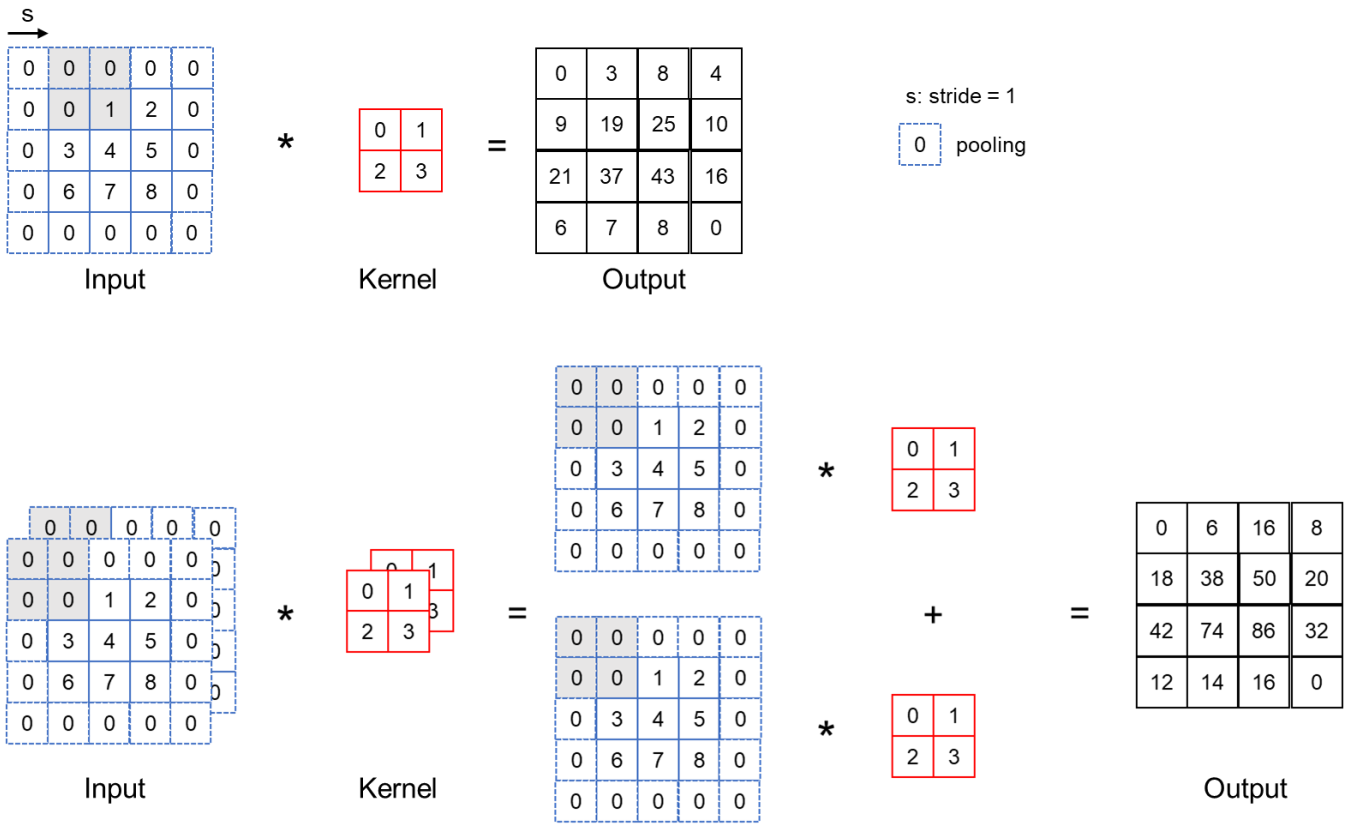

Figure 2.7: Two-dimensional convolutional neural networks operation

Recently, CNNs have remarkably improved the performance in one-dimensional spectral analysis (Zhang et al., 2019; Pullanagari et al., 2021; Martins et al., 2022). In CNNs, several convolutional layers and fully connected layers are used in the hidden layer. The convolutional layers can extract informative features based on the kernel (Figure 2.8). Normally, convolutional layers can reduce the dimension of the input features. For example, an input spectral feature with the dimension of 1×5 is transferred to a dimension of 1×2 with a kernel with a size of three and stride of two (Figure 2.8). Convolutional layers create sparse connections through a shared kernel enabling the extraction of local absorption peak features and reducing the overall number of parameters in the neural network.

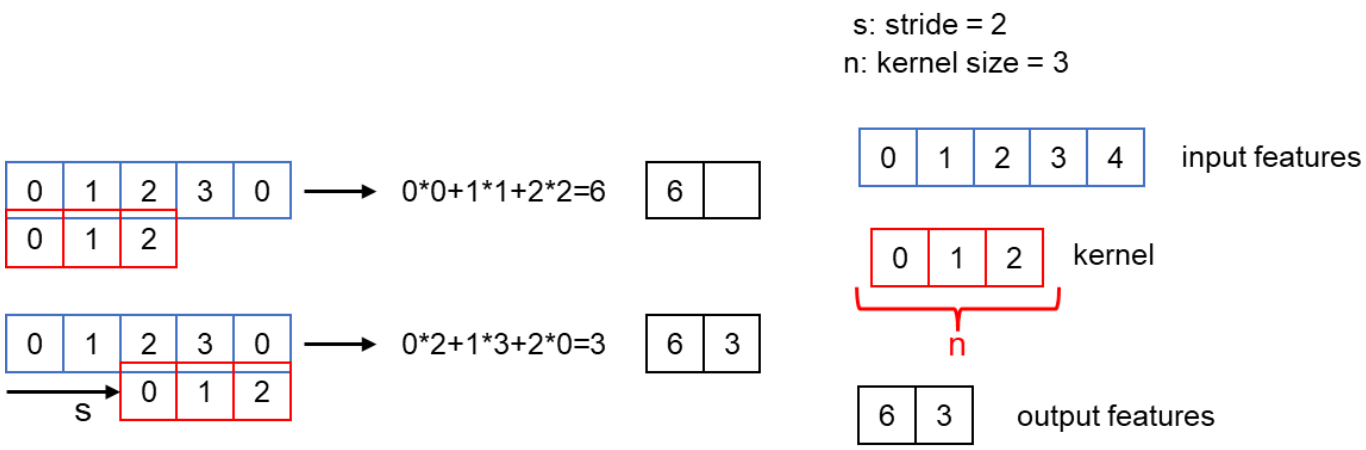

Figure 2.8: One-dimensional convolutional neural networks operation

2.5.4 Model Evaluation

A major goal of the machine learning and deep learning process is to find an algorithm that not only fits well to past data but also to future data. In order to achieve this goal, we need to split our data into training and test data sets to check the generalization ability of the algorithm. Training sets are used to train the algorithm and tune hyperparameters. Test sets are usually obtained from different locations or different seasons, which are used to assess model generalization ability. However, when there are a limited number of samples, or there are only samples from the same harvest season and origin, a typical recommendations is to spilt total samples into training and test sets using 60/40 or 70/30 ratio, through random sampling or stratified sampling (Janiesch et al., 2021). Furthermore, machine learning and deep learning algorithms have a large number of hyperparameters that can determine the model performance. One does not find the best optimal combination of hyperparameters based on the test set, as this leads to bias results. Traditionally, the training set is split into training set and validation set. the model is trained based on different combinations of hyperparameters on the training set and evaluate the performance on the validation set to choose optimal hyperparameters. The most commonly used splitting training set strategies include k-fold cross validation and bootstrapping. K-fold cross validation randomly divides the training data into k equal-sized subsets. Then the model is trained k times, each time using k-1 subsets as training data and

the remaining subset as validation data. This process ensures that each data point is used for validation exactly once. Finally, the performance metrics from each iteration are averaged to obtain a more reliable estimate of model performance. K-fold cross-validation helps in assessing the model's generalization ability and reduces the risk of overfitting or underfitting. Bootstrapping randomly samples using replacements from the training data to create multiple bootstrap samples of the same size. Each bootstrap sample is then used to compute the statistic or parameter of interest. By repeating this process numerous times, a distribution of the statistic or parameter is obtained, allowing for the estimation of its variability and confidence intervals.

The performance of machine learning and deep learning models is assessed by loss functions. Loss functions are used to measure the difference between predicted values and actual values. There are many loss functions to select from when evaluating regression or classification model performance. The commonly used loss functions in regression models include: root mean squared error (RMSE), normalized root mean squared error percentage (nRMSE%), mean absolute error (MAE) and coefficient of determinations (R^2) (Pullanagari and Li, 2021). Additionally, ratio prediction to deviation (RPD) and Lin’s concordance correlation coefficient (CCC) are also used in some spectroscopic analysis studies (Zhao et al., 2022). RPD represents the ratio of a variable's standard deviation to its standard error of prediction using a specific model. A model is considered nearly perfect when RPD exceeds 2, moderately effective when 2 > RPD > 1.4, and poor when RPD is less than 1.4 (Chang et al., 2001). The CCC, represents how closely a model predicted values confirms to the actual value along a 45-degree line from the origin. A perfect model exhibits CCC greater than 0.8, a moderate model falls within the range of 0.8 to 0.65, and a poor model is characterized by CCC less than 0.65 (Zhao et al., 2022).

For a classification model, a confusion matrix is commonly used to assess model performance. A confusion matrix provides a summary of the actual categorical levels and predicted categorical levels. Additionally, the model performance can be evaluated through

accuracy, precision, sensitivity (also known as recall), specificity and area under the curve (AUC) (Janiesch et al., 2021).

2.6 Summary and Opportunity for Research

Winemakers prefer to make wine with uniform quality of grape berries because it produces a consistent flavor and aroma, which increases the value of the wine (Bramley, 2022). Accordingly, the main objective of precision viticulture is to ensure that the quality of the harvested grapes is of the same grade. Accurate and non-destructive prediction of grape quality in both pre- and post-harvest stages can help wineries obtain grapes of the same quality grade. This review has summarized that several advanced techniques including remote sensing imagery, VNIR-SWIR spectroscopy and hyperspectral imaging system have the potential to predict the grape quality during pre- or post-harvest stage.

Predicting the quality of the grapes during the pre-harvest stage can help grape growers better specify harvest strategy to achieve selective harvesting. Remote sensing imagery can predict grape quality during the pre-harvest stage in combination with understanding grape growth conditions (Hall et al., 2011; Kasimati et al., 2021). Previous studies highlighted the possibilities of NDVI and PCD in predicting grape quality (Hall et al., 2011; Trought and Bramley, 2011; Kasimati et al., 2021). However, the correlation between other vegetation indices and grape quality can improve the prediction performance (García-Fernández et al., 2021). Additionally, the ancillary information including trunk circumference, soil electrical conductivity, weather information have also been shown to affect grape quality (Trought and Bramley, 2011; Willwerth and Reynolds, 2020). In this thesis an attempt is made to investigate the potential of using the combination of vegetation indices and ancillary variables to predict grape quality parameters.

The application of VNIR-SWIR spectroscopy in predicting wine grape quality during pre-or post-harvest stage has been well established. Whilst high prediction performance was achieved under controlled lab conditions during the post-harvest stage, there are few studies that explored the potential of applying VNIR-SWIR spectroscopy under environmental conditions

during the pre-harvest stage (dos Santos Costa et al., 2019; Kalopesa et al., 2023). Additionally, no conclusive research has been done to determine whether it is better to use spectra data in VNIR or SWIR, or VNIR-SWIR, to predict grape quality. Most studies have only split training and test sets on a single harvest season or vineyard to verify the potential of applying VNIR-SWIR spectroscopy. Gomes et al., (2017a) have suggested that the application of spectroscopic measurements of grape quality in different environmental conditions and harvest seasons are required. In order to address some of these shortcomings. This thesis explores the comparative effectiveness of VNIR, SWIR, and VNIR-SWIR spectral data in predicting various aspects of grape quality. In addition, this experiment is undertaken on different vineyards and harvest seasons to explore the practical reliability of VNIR-SWIR spectroscopy.

Although a variety of hyperspectral imaging systems have been used in research (Table 2.2). The majority of these were conducted using regression-based problem solving on image spectroscopy techniques. A strong correlation can be found between spectral data obtained from hyperspectral imaging and the quality parameters of the grape berry by developing regression methods. The primary aim of assessing grape quality revolves around distinguishing berries into two categories: ripe and unripe. Utilizing classification algorithms is important in this process of determining grape quality. With the development of deep learning algorithms such as CNNs allows the features of image data to be extracted more efficiently and rapidly. In this study, a series of deep learning algorithms are used to process the hyperspectral images and classify the grape quality parameters.

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Hongyi Lyu
Name and title of main supervisor:	Dr Miles Grafton, Senior Lecture
In which chapter is the manuscript/published work?	Chapter 3

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]
The candidate used remote sensing imagery and ancillary variables to predict grape quality based on various machine learning models.
The supervisory panel assisted with flying UAV and other technical support, editing and writing support.

Please select one of the following three options:

(●) **The manuscript/published work is published or in press**

Please provide the full reference of the research output:
Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Wei, H. E., & Sandoval, E. (2023). Using Remote and Proximal Sensing Data and Vine Vigor Parameters for Non-Destructive and Rapid Prediction of Grape Quality. Remote Sensing, 15(22), 5412. https://doi.org/10.3390/rs15225412

( ) **The manuscript is currently under review for publication**

Please provide the name of the journal:

( ) **It is intended that the manuscript will be published, but it has not yet been submitted to a journal**

Student's signature:	Hongyi Lyu 2024. 06. 29	Main supervisor's signature:	Miles Grafton Digitally signed by: Miles Grafton DN: CN = Miles Grafton email = m.grafton@massey.ac.nz C = NZ O = Massey University OU = SAE Date: 2024.06.29 13:54:30 +1200'

This form should be placed at the beginning of each relevant thesis chapter.

[1] Refer to the Massey University Publishing and Authorship guidelines (OneMassey for staff, Stream for students) and/ or Contributor Roles Taxonomy (CRediT) guidelines for guidance.

3 Using Remote and Proximal Sensing Data and Vine Vigor Parameters for Non-Destructive and Rapid Prediction of Grape Quality

Remote sensing imagery can predict grape quality during the pre-harvest stage by understanding grape growth conditions. There is a correlation between vegetation indices and grape quality, that can improve the prediction performance of algorithms. Additionally, the ancillary information such as trunk circumference, soil electrical conductivity and topography has also been shown to affect grape quality. This chapter is dedicated to investigating the potential of using remote sensing imagery and ancillary variables to predict the grape quality during the pre-harvest stage. This can be revealed using various machine learning algorithms with an objective of evaluating different machine learning algorithms prediction performance. This is an important step towards achieving selective harvesting of berries of different quality.

This chapter based on:

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Wei, H. E., & Sandoval, E. (2023). Using Remote and Proximal Sensing Data and Vine Vigor Parameters for Non-Destructive and Rapid Prediction of Grape Quality. *Remote Sensing*, *15*(22), 5412. https://doi.org/10.3390/rs15225412

Abstract

The traditional method for determining wine grape total soluble solid (TSS) is destructive laboratory analysis, which is time consuming and expensive. In this study, we explore the potential of using different predictor variables from various advanced techniques to predict the grape TSS in a non-destructive and rapid way. Calculating Pearson's correlation coefficient between the vegetation indices (VIs) obtained from unmanned aerial vehicles (UAV) multispectral imagery and grape TSS resulted in a strong correlation between OSAVI and grape TSS with a coefficient of 0.64. Additionally, seven machine learning models including ridge regression and lasso regression, k-Nearest neighbor (KNN), support vector regression (SVR), random forest regression (RFR), extreme gradient boosting (XGBoost), and artificial neural network (ANN) are used to build the prediction models. The predictor variables include the UA derived VIs, and other ancillary variables including normalized difference vegetation index (NDVI_proximal) and soil electrical conductivity (EC_a) measured by proximal sensors, elevation, slope, trunk circumference, and day of the year for each sampling date. When using 23 VIs and other ancillary variables as input variables, the results show that ensemble learning models (RFR, and XGBoost) outperform other regression models when predicting grape TSS, with the average of root mean square error (RMSE) of 1.19 and 1.2 °Brix, and coefficient of determination (R^2) of 0.52 and 0.52, respectively, during the 20 times testing process. In addition, this study examines the prediction performance of using optimized soil adjusted vegetation index (OSAVI) or normalized green-blue difference index (NGBDI) as the main input for different machine learning models with other ancillary variables. When using OSAVI-based models, the best prediction model is RFR with an average R^2 of 0.51 and RMSE of 1.19 °Brix, respectively. For NGBDI-based model, the RFR model showed the best average result of predicting TSS were a R^2 of 0.54 and a RMSE of 1.16 °Brix, respectively. The approach proposed in this study provides an opportunity to grape growers to estimate the whole vineyard grape TSS in a non-destructive way.

Keywords: wine grape; vegetation indices; UAV multispectral imagery; sugar content

3.1 Introduction

Viticulture is an important sector within the New Zealand horticultural system, with a total vineyard production area extending over 41,000 ha (New Zealand Winegrowers, 2023). In New Zealand, conventional vineyards typically employ uniform management practices. However, studies have demonstrated the spatial and temporal variation in vine vigor, grape yield and quality within the vineyard scale (Baluja et al., 2013; Bramley et al., 2011b). With uniform management a single application rate of fertilizers, pesticides and irrigation are used for the entire vineyard. As a consequence of uniform management, parts of the field are likely to receive too little and others too much input. This could have negative impacts on the environment such as groundwater pollution, soil degradation and increase pressure from weeds and pests (Froment et al., 1995). Precision viticulture (PV) provides an opportunity for grape growers to understand and manage the vineyard spatial variability by using remote sensing data and geostatistical analysis. Remote sensing allows viticulturists to continuously monitor spatial and temporal variation in soil properties and vine growth status (Wei et al., 2023). This is particularly relevant for New Zealand wine growers who need to understand and manage vineyard variability to increase grape quality and achieve sustainable viticulture.

Remote sensing is the process of acquiring spectral data remotely from several platforms including ground vehicles, aircraft, uncrewed aerial vehicles (UAV), and satellites. In recent years, remote sensing has been widely applied in PV to evaluate the vineyard spatial variability of vine vigor, nutrient status, wine water status, grape yield and quality (Lamb et al., 2004; Rey-Caramés et al., 2015; Wei et al., 2023). Satellite platforms with multispectral cameras such as Sentinel-2 with 10 m pixel resolution and Landsat 8 with 30 m pixel resolution have been used on a regional scale to predict grape yield (Arab et al., 2021; Arab and Ahamed, 2022). However, low spatial resolution satellite imagery makes it difficult to monitor vine growth status on a vineyard block scale without bias, as a single pixel of satellite imagery often mixes inter-row crops and bare soil. UAV and aircraft sensors carrying multispectral or hyperspectral cameras have provided high spatial resolution imagery on a vineyard scale. For example, Carrillo et al.

found a linear relationship between berry weight and NDVI derived from multispectral airborne imagery (Carrillo et al., 2016). In addition, Lamb et al. found a strong correlation between berry quality parameters and multispectral airborne imagery obtained at veraison (Lamb et al., 2004). However, the operational cost of aircraft is very expensive. Compared to other platforms, UAV provide an interesting alternative approach for PV, as it can provide high spatial resolution images at a lower cost (Matese et al., 2015). For example, García-Fernández et al. found a significant correlation between RGB-based vegetation indices derived from UAV imagery and berry quality (García-Fernández et al., 2021). Wei et al., (2022) found the VIs derived UAV multispectral imagery combined with other environmental variables can predict grapevine water status (stem water potential Ψstem), with the RMSE of 138 kPa. In addition to remote sensing, proximal sensors have demonstrated their capability to explore the spatial and temporal variation of vine growth status within the vineyard scale. Bramley et al. used a handheld proximal sensor (Crop Circle™) and an EM38 electromagnetic soil sensor to explore the spatial variability of vine vigor and yield as well as soil texture on a vineyard block scale, while neither EC_a, nor VIs derived from proximal sensor were good predictors of grape yield (Bramley et al., 2011b). Their result also indicated that EC_a and VIs significantly correlated with vine vigor. Furthermore, a portable hyperspectral spectroradiometer can be used to predicted the vine growth status and grape quality within the vineyard scale (Kalopesa et al., 2023; Lyu et al., 2023a).

One of the main objectives of PV is selective harvesting. A major reason for wishing to do this, considering the variability of vineyards, is to harvest berries with consistent quality during the harvest stage resulting in higher profit margins from wine (Bramley et al., 2003). In New Zealand, sugar content commonly describes the quality of wine grapes at harvest. The sugar content relates to the alcohol level of wine during the fermentation process. The traditional method used in monitoring wine grape sugar content is to perform sample-based laboratory destructive analysis which can be time consuming and an expensive process. In recent years, remote and proximal sensors have been used to monitor grape sugar content (Benelli et al., 2021; García-Fernández et al., 2021; Kasimati et al., 2021). For example, Benelli et al. used a

push broom hyperspectral camera mounted on a vehicle to predict the sugar content of wine grapes (Benelli et al., 2021). In addition, Kasimati et al. used NDVI obtained from proximal and remote sensing during different growing stages to predict grape sugar content with R^2 values of 0.61. Presently, the increase in computing power and advanced sensing techniques enable more accurate prediction of grape quality, helping grape growers assess grape quality before harvest and thus develop a selective harvesting plan.

Previous research has been conducted using Pearson's correlation coefficient and machine learning techniques to explore the relationship between spectral indices data and vine growth status, yield, and berry quality parameters. Pearson's correlation coefficients have been constructed to select key spectra indices relevant to the grape sugar content (García-Fernández et al., 2021). Machine learning techniques have been constructed to model both linear and non-linear relationships between spectral indices data and vine growth status, yield, and berry quality. One study used an artificial neural network (ANN) to predict the table grape yield through different vegetation indices (VIs) obtained from satellite remote sensing (Arab et al., 2021). In regard to studies conducted in predicting wine grape berry quality parameters, many studies used hyperspectral imaging systems to predict various quality parameters such as sugar content, pH and titratable acidity through different machine learning techniques (Dambergs et al., 2015). However, most of these studies used remote or proximal sensing to directly measure berries or clusters. Few studies have predicted grape quality through the VIs derived from canopy or leaf level. For example, Kasimati et al. achieved good prediction accuracy with R^2 values of 0.65 for grape sugar content using NDVI and automated machine learning (Kasimati et al., 2022). The VIs measured from vine canopy or leaves can reflect the vine vigor, water status and nutrient status (Bramley, 2022; Lyu et al., 2023a; Wei et al., 2023). Additionally, the ancillary information including trunk circumference, soil electrical conductivity, weather information have also been shown to affect grape quality (Trought and Bramley, 2011; Willwerth and Reynolds, 2020). Thus, it is important to explore the possibility of using the combination of VIs, ancillary information, and machine learning techniques to predict grape quality parameters.

The aims of this study are to (1) map spatial variability of vine growth status based on remote sensing imagery, (2) to create an alternative method to predict grape sugar content thorough VIs, derived from proximal and remote sensing, and other ancillary environmental variables during different growth stages, and (3) to determine the ability of different machine learning techniques in predicting grape sugar content.

3.2 Methods

3.2.1 Study Sites

The study research sites were conducted in two commercial wine grape vineyard blocks on the Palliser Estate located in Martinborough, New Zealand (175.45235E 41.21119°S, WGS 84). The study vineyards are named Hua Nui (HN) and Pencarrow (PN). The variety used in the study areas was Pinot Noir which were planted in 1998-2000 for winemaking. The research sites selected for data collection were 3.31 ha for HN and 7.51 ha for PN. The wine grapes in study sites were trained with two-cane vertical shoot positioning. Inter- and intra-row planting space is 2.2 and 1.7m for HN, and 2.2 and 1.8m for PN. The region of Martinborough has a mild coastal climate with an average temperature of 18 °C. The soils in the vineyards are mostly clay and silty loams, which are known to have moderate soil water holding capacity. Vine phenology in research vineyards is shown in Table 3.1. The vineyard manager is responsible for managing the vines and applied all inputs.

Table 3.1: Vine phenology in the study vineyards

Vine phenology	Date
Budburst	September, October
Leaf development	October, November
Inflorescence emergence	November
Flowering	November, Early December

Fruit set and fruiting	December, January
Veraison	Late January, February
Harvest	March, April

3.2.2 Grape Sugar Content Data Acquisition

Wine grapes were manually harvested between February 20th and March 14th, 2023 (from veraison to harvest time). Figure 3.1 show the sampling locations in each vineyard. The sampling locations were collected from a predetermined nested grid. Each grid had locations taken at intervals of 4m, 8m, 16m and 32m. Sampling locations were mapped prior to field data collection. During each measurement, the healthy vine closest to the sampling location was selected for sampling. The location of the sampling vine's trunk was recorded by a global navigation satellite system (GNSS) with real-time kinematic (RTK) correction (model: GPS1200+, Leica Geosystems AG., Heerbrugg, Switzerland). The weather during collecting samples period was shown in Figure 3.2. Three berries from a single bunch were randomly selected from each sampling vine. Table 3.2 show the number of sampling vines from each measurement date. After acquiring grape berries, total soluble solids (TSS) expressed in °Brix, was chosen as a proxy for grape sugar content and measured by using a portable digital (PAL-ALPHA Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan). The sugar refractometer content of each sampling vine was determined by calculating the mean of the three measurements taken per sampled vine.

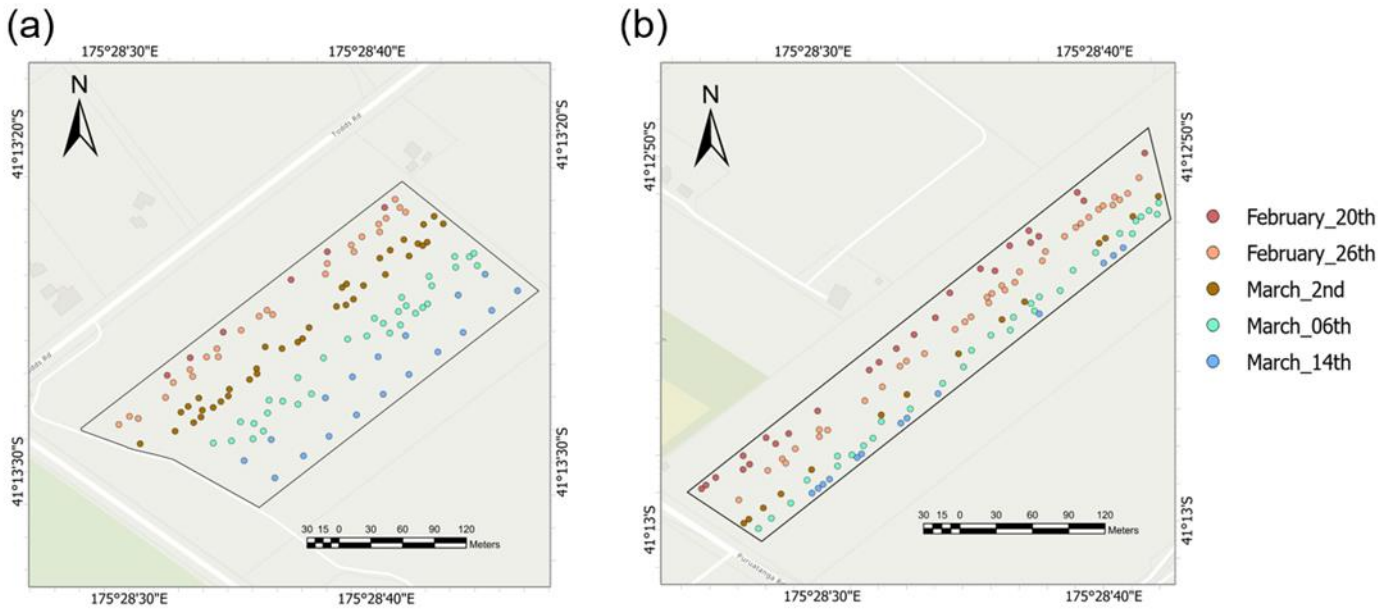

Figure 3.1: The location of sampling vines in PN (a); HN (b) (Points represent the location of sampling vines)

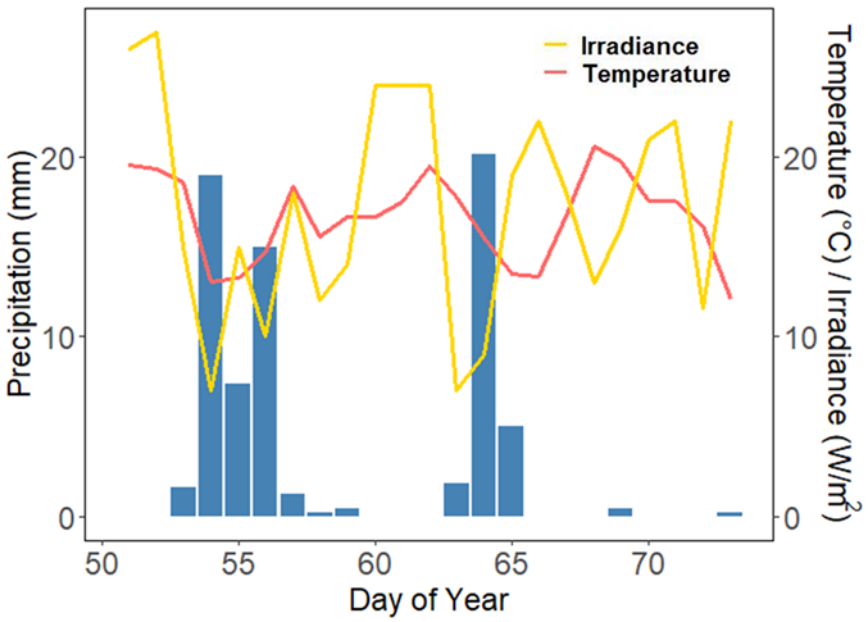

Figure 3.2: Total daily precipitation, average temperature and irradiance recorded by on-site weather station

Table 3.2: The number of sampling vines in each measurement day

Vineyard	20th February	26th February	02nd March	06th March	14th March
Pencarrow	6	25	37	32	18
Hua Nui	25	36	14	30	13

3.2.3 Canopy and Leaf Reflectance Data Acquisition

Multispectral imagery was acquired by DJI P4 Multispectral (Da-Jiang Innovations, Shenzhen, China) between 11.00 and 14.00 under stable weather conditions and a clear sky on the 14th December 2022 (flowering stage). The DJI P4 Multispectral was gimbal equipped, with a 6-camera array, one camera is an RGB camera, designed for visible light and capturing standard photos, while the remaining five cameras are dedicated to capturing 2MP images, at various wavelengths: near-infrared, red-edge, red, green, and blue, with center wavelengths at 840nm, 730nm, 650nm, 560nm and 450nm, and bandwidth of ±26nm, ±16nm, ±16nm, ±16nm and ±16nm respectively. The DJI P4 flew at 80 m height capturing images with a spatial resolution of 4.2 cm for PN, and some at 84.9m resulting in 4.5 cm spatial resolution imagery for HN. The DJI P4 is equipped with an integrated sunlight sensor that records irradiance during flight. This sensor captures data in the same bands as the multispectral sensor, which is used for radiometric calibration purposes. In addition, radiometric calibration was performed on the image blocks, using reference images acquired from a calibrated reflectance panel (Seattle, WA, USA). To assure the accuracy of the image, several ground control points were recorded by GNSS-RTK correction in each vineyard during flight for geometric correction. Photogrammetric processing was applied to the georeferenced multispectral im-ages using Pix4Dmapper (Pix4D SA, Lausanne, Switzerland). The digital surface models (DSM), digital terrain models (DTM), and reflectance maps of the study sites were created by Pix4Dmapper.

Due to discontinuous vegetation surfaces for the vineyard features, it becomes essential to differentiate between the pixels belonging to the vine canopy and those in the inter-row spacing. The process of detecting vine rows consists of the following steps: (1) NDVI were calculated by band math function in ENVI 5.6 (Research Systems Inc., Boulder, Co, USA), producing a greyscale image for each study site; (2) then, a global threshold was implemented on the NDVI images to generate a binary image based on Otsu's method (Otsu threshold value is 0.535); (3) a single greyscale image was calculated from subtraction of the DTM to the DSM for each study site; (4) then, a height threshold of 0.9 m was implemented on the greyscale images to generate a binary image (Wei et al., 2022). The height threshold was dependent on the vineyards' architecture; (5) then these two binary images were converted to a feature map by ArcGIS Pro 2.9 (ESRI, Redlands, CA, USA); (6) then, using the "Clip" tool in ArcGIS pro to extract the overlapping portion of these two features; (7) using the vine row feature map as a mask layer to retain the original multispectral images (Figure 3.3).

We located the image pixels that corresponded to the 236 sampling points by selecting pixels on both sides of each sampling point, considering the vine spacing within each vineyard. From these pixels, we extracted reflectance values, averaged them, and subsequently calculated the VIs using the formulas outlined in Table 3.3.

Leaf reflectance data was measured during 17th February 2023 (veraison) to record the NDVI of the grapevines' leaves. Three leaves in each sampling vine were randomly selected and scanned using a handheld RapidSCAN CS-45 (Holland Scientific Inc., Lincoln, NE, United States) to record these three leaves NDVI value (NDVI_proximal). The average of the three leaves NDVI values (NDVI_proximal) was used to represent the sampling vine growth status during veraison.

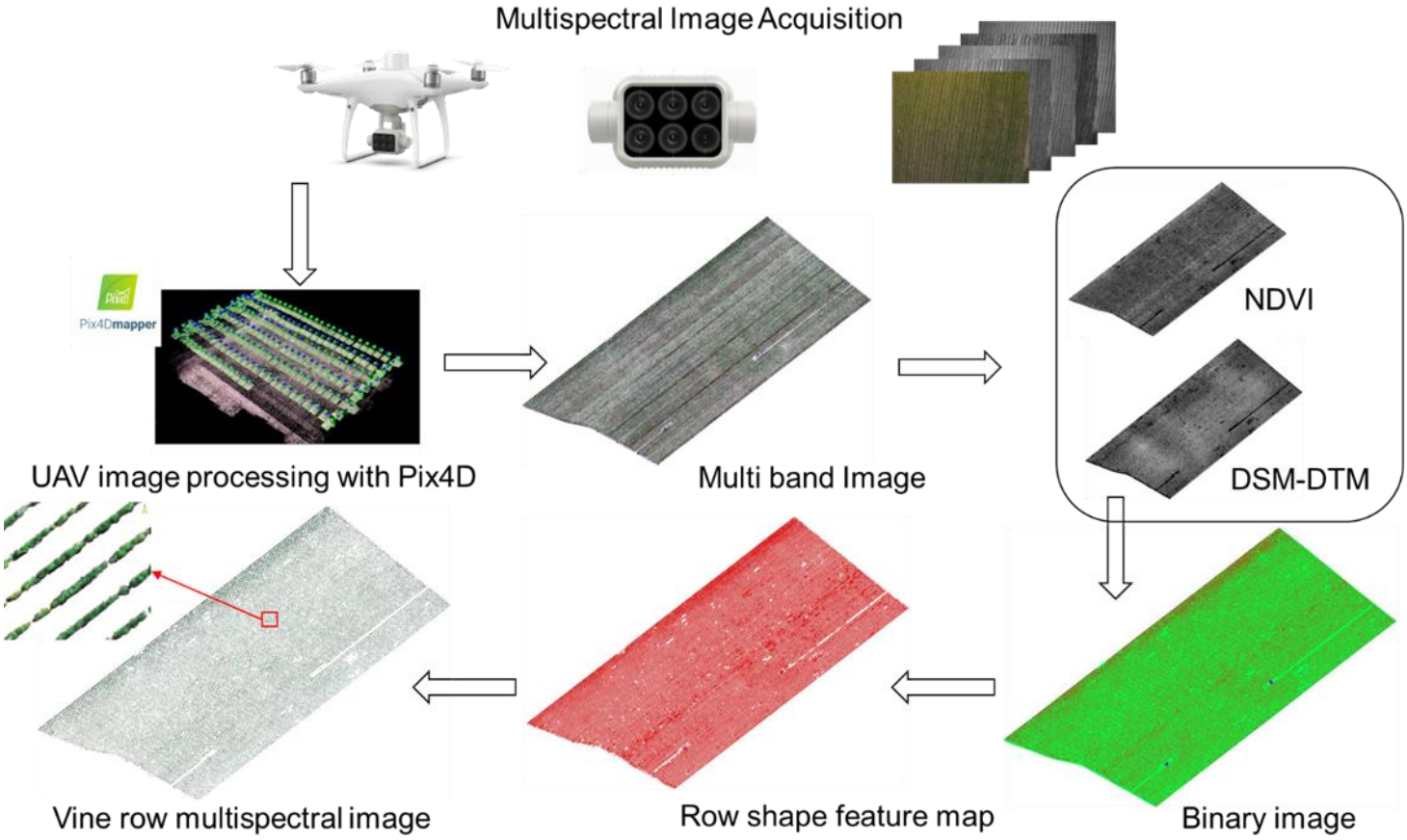

Figure 3.3: Vine row segmentation workflow

Table 3.3: The spectral indices used in this study

Vegetation indices	Formula	Reference
PCD	NIR/Red	(Arnó Satorra et al., 2009)
NDVI	$(NIR - Red)/(NIR + Red)$	(Rouse Jr et al., 1974)
GNDVI	$(NIR - Green)/(NIR + Green)$	(Romero et al., 2018)

MSR	$((\frac{NIR}{Red}) - 1)/\sqrt{((\frac{NIR}{Red}) + 1)}$	(Gil-Pérez et al., 2010)
MSAVI	$\frac{2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - Red)}}{2}$	(Gil-Pérez et al., 2010)
RDVI	$(NIR - Red)/\sqrt{(NIR + Red)}$	(Gil-Pérez et al., 2010)
OSAVI	$(1 + 0.16)\ \frac{NIR - Red}{NIR + Red + 0.16}$	(Gil-Pérez et al., 2010)
R/B index	$Red/Blue$	(Jiménez-Brenes et al., 2019)
R/G index	$Red/Green$	(Jiménez-Brenes et al., 2019)
NGRDI	$(Green - Red)/(Green + Red)$	(Jiménez-Brenes et al., 2019)
NPCI	$(Blue - Red)/(Blue + Red)$	(Jiménez-Brenes et al., 2019)
VARI	$(Green - Red)/(Green + Red - Blue)$	(Brook et al., 2020)
Clgreen	$NIR/Green - 1$	(Jiménez-Brenes et al., 2019)
ARI	$Green^{-1} - Rededge^{-1}$	(Albetis et al., 2017)
MARI	$(Green^{-1} - Rededge^{-1}) * NIR$	(Albetis et al., 2017)
CLREDEDGE	$(NIR/Rededge) - 1$	(Soubry et al., 2017)
MCARI	$\big((Rededge - Red) - 0.2 * (Rededge - Green)\big) * (Rededge/Red)$	(Soubry et al., 2017)
NDRE	$(NIR - Rededge)/(NIR + Rededge)$	(Soubry et al., 2017)
EVI	$2.5 * (NIR - Red)/(NIR + 6 * Red - 7.5 * Blue + 1)$	(Cogato et al., 2019)

NGBDI	$(Green - Blue)/(Green + Blue)$	(Pádua et al., 2018)
G%	$Green/(Red + Green + Blue)$	(Pádua et al., 2018)
REGI	$(Rededge - Green)/(Rededge + Green)$	(Albetis et al., 2018)
RERI	$(Rededge - Red)/(Rededge + Red)$	(Albetis et al., 2018)

3.2.4 Vine Vigor Parameter Acquisition

The trunk circumference (TC) was chosen as a proxy for vine vigor. The trunk circumference of all of the vines was measured 10 cm above the graft union and 10 cm below the head of the vine with the paper ruler during the bud and leaf growing stage. These two measurements per sampled vine were averaged to represent the vine vigor (Bramley et al., 2011a).

3.2.5 Soil and Terrain Data Acquisition

Many studies indicated that soil electrical conductivity (EC_a) can be used to assess soil texture types and water content (SU et al., 2014; Trought et al., 2008). Soil EC_a is widely used to explore the spatial variation in soil properties within the vineyards (Bramley et al., 2011a). In addition, one study showed that soil EC_a is directly related to vine water status, berry weight and sugar content (Yu et al., 2020). In this study, an electromagnetic induction sensor EM38-MK2 (Geonics Ltd., Mississauga, ON, Canada) was used to assess the soil EC_a in the study area during 27th May 2021. The EM38-MK2 was operated in the vertical dipole mode, capturing integrated EC_a measurements at a depth of approximately 1.5 m. The EM38-MK2 was towed at the back of an all-terrain vehicle, maintaining a distance of less than 0.2 m between the instrument and the vehicle. To ensure accurate georeferencing of all point data from the EC_a (mS/m) survey, a Trimble Yuma tablet equipped with an onboard GPS receiver (model: Yuma, Trimble) accurate to 2-4 m, was utilized. Soil EC_a points were measured at intervals of approximately 3-10 m along transects, with a 10 m spacing between individual measurements.

The elevation (m) and slope (degree) data for the study site were obtained from the 'Wellington LiDAR 1m DEM (2013-2014)' layer. This dataset was made accessible through the Land Information New Zealand data service (https://data.linz.govt.nz/, accessed on 28th July 2023). This digital elevation model (DEM) has a resolution of 1 m and was created using aerial LiDAR data captured between 2013 and 2014, encompassing the study vineyards.

3.2.6 Geostatistical Analysis

The application of geostatistics in PV has been undertaken in many studies (Bramley et al., 2011b). The purpose of geostatistics is describing spatial autocorrelation of a regionalized variable (the georeferenced data) and uses this information to predict the values of the variable across an entire field by Kriging. The central tool in geostatistics is the variogram, which is a set of semi-variances plotted against the lag distances between the measurements to describe the way in which a property varies from place to place. Experimental variograms were computed using R statistical software (R Core Team, version 4.2.2) with the package “gstat” by the Matheron’s method of moments, as in the following formula:

$$\gamma(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} [Z_x - Z_{(x+h)}]^2 \qquad \text{Equation 3-1}$$

Where $2m(h)$ is the number of paired comparisons at the lag interval h, $Z_x$and $Z_{(x+h)}$ are the values of the property at two locations separated by distance h.

After computing the experimental variogram, model the experimental variogram based on a suitable mathematical model. The best fitting mathematical model was selected based on the lowest residual sum of squares (RSS). The parameters of the fitted variogram were used to interpolate the trunk circumference, NDVI, PCD, soil EC_a, slope and elevation value based on ordinary Kriging by using ArcGIS Pro 2.9 (ESRI, Redlands, CA, USA). The soil EC_a values less than 0 mS/m were removed before doing geostatistical analysis. The Kriging interpolation images were then exported to a raster layer with the same gride size as used by the multispectral UAV image. For each sampling vine, the mean values of NDVI, PCD, soil EC_a, elevation and slope were computed using “zonal statistics as table” in ArcGIS Pro 2.9.

3.2.7 Machine Learning Model

Different machine learning models were performed to predict the sugar content. The input variables include, canopy reflectance data, leaf reflectance data, TC, soil EC_a, elevation, slope, and day of year for sampling date (DOY). The machine learning models used; included regularized regression, k-Nearest Neighbors (KNN), support vector regression (SVR), random forest regression (RFR), XGBoost and ANN.

Regularized regression: regularized regression is used to explore the linear relationship between input and output variable, when the data set contain more features than observations. In addition, it is suitable for the analysis of multicollinearity among the features. The objective function of a regularized regression is in the following formula:

$$minimize\ (\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + P) \qquad \text{Equation 3-2}$$

Where n is the sample number, y_i is the measured truth value of the ith sample, $\hat{y}_i$ is the predicted value of the ith sample, P is the penalty term.

In this study we used ridge penalty and lasso penalty to build the linear relationship between input and output variables. When using the ridge penalty, the formula of penalty parameters is

$$P = \lambda \sum_{j=1}^{p} \beta_j^2 \qquad \text{Equation 3-3}$$

When using the lasso penalty, the formula of penalty parameters is

$$P = \lambda \sum_{j=1}^{p} |\beta_j| \qquad \text{Equation 3-4}$$

Where $\lambda \geq 0$ is a tuning parameter, p is the feature number, β_j is the regression coefficient of the jth feature.

KNN: It aims to identify the k most similar instances from a training dataset to predict the target value of a new data point. The Euclidean distance metric is commonly used to determine the similarity between observations. The formula of Euclidean distance is

$$\sqrt{\sum_{j=1}^{p}(x_{aj} - x_{bj})^2}$$ Equation 3-5

Where x_a and x_b represent the observations, j represents the feature, p represents the feature number.

SVR: Support vector machine (SVM) try to find a hyperplane in an N-dimensional space (N-the number of features) that “best” classify the two classes. Hyperplane represents a decision boundary that help classify the data points. SVR is an extend tool of SVM to solve regression problems.

RFR: It is an ensemble learning model which combines the multiple decision trees on different subsets of the training data (bootstrap samples), and then averaging their predictions to achieve more accurate and robust regression results.

XGBoost: It combines gradient boosting with regularization techniques, creating a boosted ensemble of decision trees. XGBoost optimizes model predictions by fitting negative gradients of the loss function during each boosting iteration.

ANN: It consists of interconnected nodes called neurons organized into layers: an input layer, one or more hidden layers, and an output layer. Each neuron receives input data, performs computations, and passes the results to the next layer through weighted connections. Activation functions introduce non-linearity and enable ANN to learn complex relationships in data. In this study we use a single layer neural network to predict the TSS based on input variables.

To evaluate the model’s prediction performance, the dataset is divided into training and test sets with the ratio of 7:3. This process was repeated 20 times with different data splits to improve the estimated performance of study models. The performance of machine learning models is affected by their hyperparameters. Thus, it was important to tune the hyperparameters. Bayesian hyperparameter optimization was used on the training set with 10-fold cross-validation to search for the best combination of hyperparameters based on the root mean square error (RMSE). Appendix 1 show different regression models’ hyperparameter

settings. In addition, the coefficient of determination (R^2) and RMSE were selected to compare the performance of different machine learning models on the test set. The Waller-Duncan test was used to conduct multiple comparisons between different machine learning models. The formula of R^2 and RMSE was as followed:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{Equation 3-6}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{Equation 3-7}$$

where n is the number of sampling vines, y_i is the measured TSS value of the ith vine sample, $\bar{y}$ is the mean measured TSS value, $\hat{y}_i$ is the model predicted TSS value of the ith vine sample.

3.3 Result

3.3.1 Variation in Total Soluble Solids

Both vineyards were measured five times from veraison in late February to harvest in middle March. Measuring TSS in this period is crucial for assessing grape maturity and consequently, making informed decisions about selecting the optimal harvest day. Figure 3.4 showed the variation in TSS collected from 236 vines across two distinct vineyards, along with the distribution of these measurements for each sampling date. The grape TSS values ranged from 12.8 to 21.2 °Brix during the study period. The extensive range of data within the dataset facilitated the construction of a robust calibration model. The grape TSS values increased rapidly from 02nd to 14th March suggesting a gradual progression towards ripeness as TSS accumulate. It is worth noting that the TSS measured on 20th and 26th February were higher compared to those of 2nd March. One possible explanation is the variability of sample collection locations lead to spatial variation in grape TSS. In addition, the height of the boxplot and the magnitude of the interquartile range represent the spatial variability of TSS within each vineyard during different measurement dates (Figure 3.4).

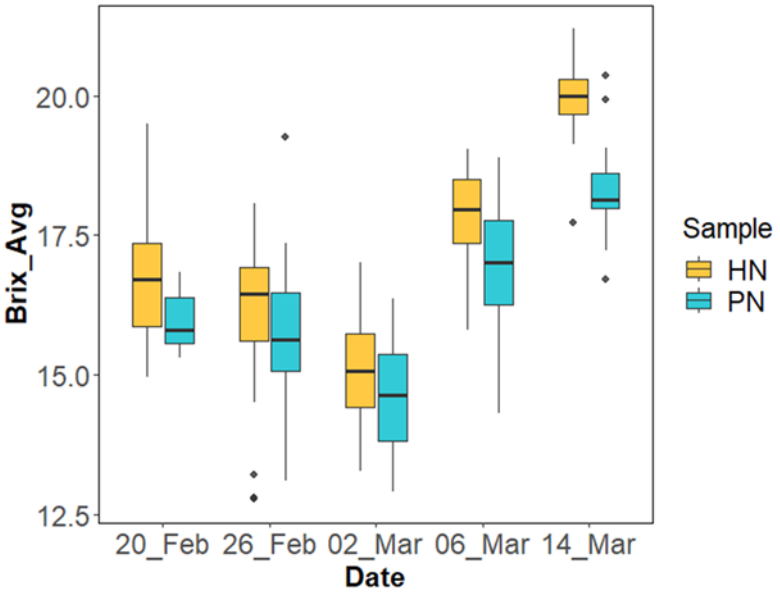

Figure 3.4: The boxplot of grape TSS during study period

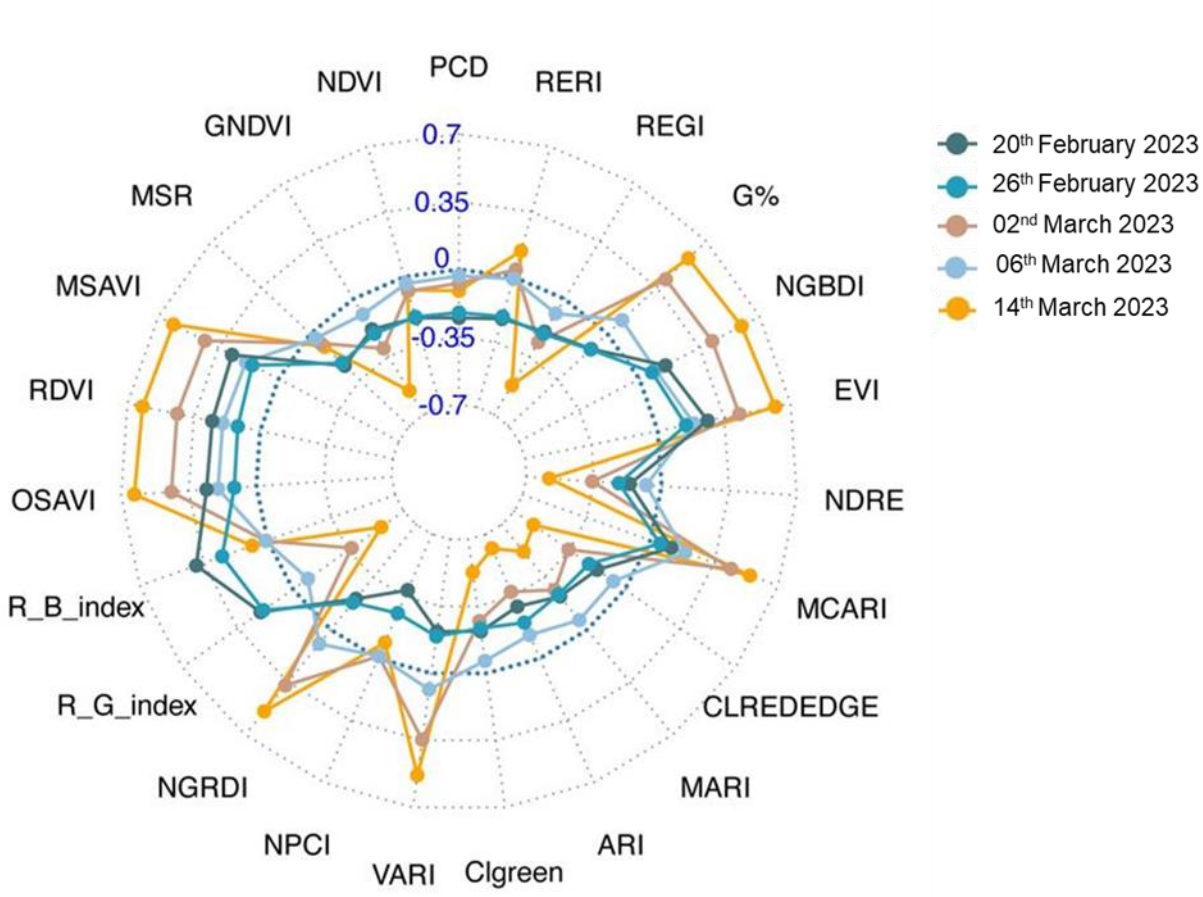

Figure 3.5: Pearson's correlation coefficient between VIs and grape TSS (Different colors represent different sampling date)

3.3.2 Pearson's Correlation Coefficient between VIs and TSS

In this study, 23 VIs frequently used in viticulture were calculated based on the red, red edge, green, blue and near infrared (NIR) bands of the UAV multispectral images. The Pearson's correlation coefficient was used to explore the relationship between VIs and grape TSS, to check the potential of using UAV multispectral images to predict the grape quality. In order to mitigate the influence of different sampling dates on grape TSS content, we calculated the Pearson's correlation coefficient between VIs and grape TSS for each respective sampling date (Figure 3.5). When the grapes reached the harvest stage, most of VIs showed a strong correlation with grape TSS with the absolute value of the Pearson's correlation coefficient, greater than 0.5. Among all the VIs, optimized soil adjusted vegetation index (OSAVI) had the maximum correlation with grape TSS, with a coefficient of 0.64 during the harvest stage. This

was followed by renormalized difference vegetation index (RDVI), enhanced vegetation index (EVI), and anthrocyanin reflectance index (ARI) with the absolute value of the Pearson's coefficient greater than 0.62 during the harvest stage. The NDVI and PCD which are widely used in assess vine vigor growth status show a negative correlation with grape TSS from veraison to harvest stage. This is because vine vigor has an negative impact on grape quality (Bramley et al., 2011a). However, most of VIs show a weak correlation with grape TSS during veraison stage. In addition, the simple linear relationship between each 23 VIs and grape TSS were explored. The result showed that OSAVI had the best linear correlation with grape TSS with a R^2 of 0.4 and RMSE of 0.89 (Table 3.4). This was followed by RDVI, ARI, EVI, MSAVI, and NGBDI (Table 3.4). The OSAVI is calculated based on NIR and red bands, this suggests that OSAVI is a promising candidate to predict the grape TSS when multispectral sensors are available. The promising performance of NGBDI implies that, when only RGB sensors are available, NGBDI may serve as a spectral indicator for grape TSS. In practical applications, RGB sensors are less expensive than multispectral sensors. Thus, it is important to explore which sensor is better for predicting grape quality.

Table 3.4: R^2 and RMSE based on a linear regression between TSS and each VIs during March 14th

VIs	R^2	RMSE
OSAVI	0.40	0.89
RDVI	0.38	0.90
ARI	0.38	0.90
EVI	0.38	0.90
MSAVI	0.38	0.90

NGBDI	0.35	0.92
NDRE	0.34	0.93
CLREDEDGE	0.33	0.93
G%	0.33	0.94
R_G_index	0.31	0.95
GNDVI	0.31	0.95
MCARI	0.30	0.96
NGRDI	0.30	0.96
VARI	0.28	0.97
Clgreen	0.28	0.97
REGI	0.27	0.97
MARI	0.27	0.98
RERI	0.02	1.13
PCD	0.01	1.14
MSR	0.01	1.14
R_B_index	0.01	1.14
NDVI	0.01	1.14
NPCI	0.01	1.14

3.3.3 Spatial Variability of Soil EC_a, Elevation and Vine Vigor Status

The geostatistical analysis was used to map the spatial variation in soil EC_a, elevation and vine vigor status within the vineyard. Soil EC_a was measured at 0.5m depth by an EM38-MK2. The kriging interpolation maps showed the spatial variability of soil EC_a, elevation and vine vigor status within the vineyard (Figure 3.6, 3.7). In the PN vineyard, soil EC_a values showed lower values in the northeastern portion, higher values in a small portion of the eastern region, and the south-western border (Figure 3.6a). In the HN vineyard, soil EC_a values were low in the south-western section of the vineyard, and in the north-eastern section showed high soil EC_a values (Figure 3.7a). When it turns to elevation, elevation values were high in the south-eastern region and low in the north-western region in PN (Figure 3.6b). In the HN vineyard, elevation values were higher in the middle region and lower in the north-eastern region. The trunk circumference, NDVI and PCD were used to represent the vine vigor status in many studies (Bramley, 2022). Figures 5c and 6c show the spatial variation in TC within study sites. In the PN vineyard, TC values were lower in the south-eastern boundary region and higher in the north-western region (Figure 3.6c). In HN vineyard, TC values were low in the north-eastern boundary region and high in the northern region (Figure 3.7c). When it turns to NDVI and PCD, the spatial pattern of these two VIs is similar in each vineyard (Figure 3.6, 3.7). In PN, NDVI and PCD showed high value in the south-western and north-eastern corner, while the low value

show in middle region (Figure 3.6d, e). In HN, NDVI and PCD showed high value in the south-eastern boundary, and low value in the northern region. (Figure 3.7d, e).

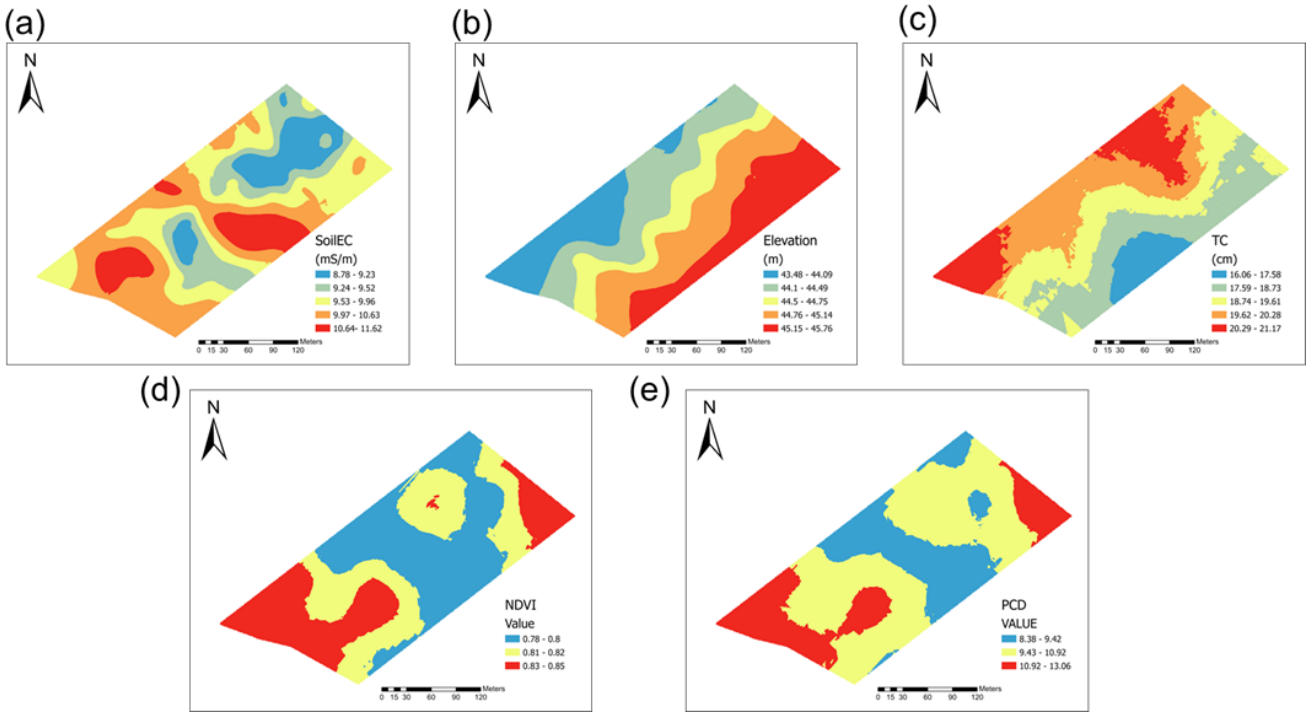

Figure 3.6: Spatial interpolation map of soil EC_a (a); elevation (b); trunk circumference (c); NDVI (d); PCD (e) in PN vineyard

3.3.4 Prediction Model Performance of Grape TSS

The machine learning models used both linear and nonlinear regression analysis to predict grape TSS. The input variable includes VIs obtained from UAV imagery, NDVI_proximal and soil EC_a obtained from proximal sensor, elevation and slope obtained from a LiDAR camera, as well as TC obtained directly in the vineyards. In order to ensure the robustness and generalization of the models used in this study, the model evaluation process was repeated 20 times with different data splits. When using linear regression models, the best prediction model is ridge regression with an average R^2 of 0.31 and RMSE of 1.43 °Brix, respectively (Table 3.5). For lasso regression, the average result of predicting TSS were a R^2 of 0.3 and a RMSE of 1.44 °Brix respectively (Table 3.5)

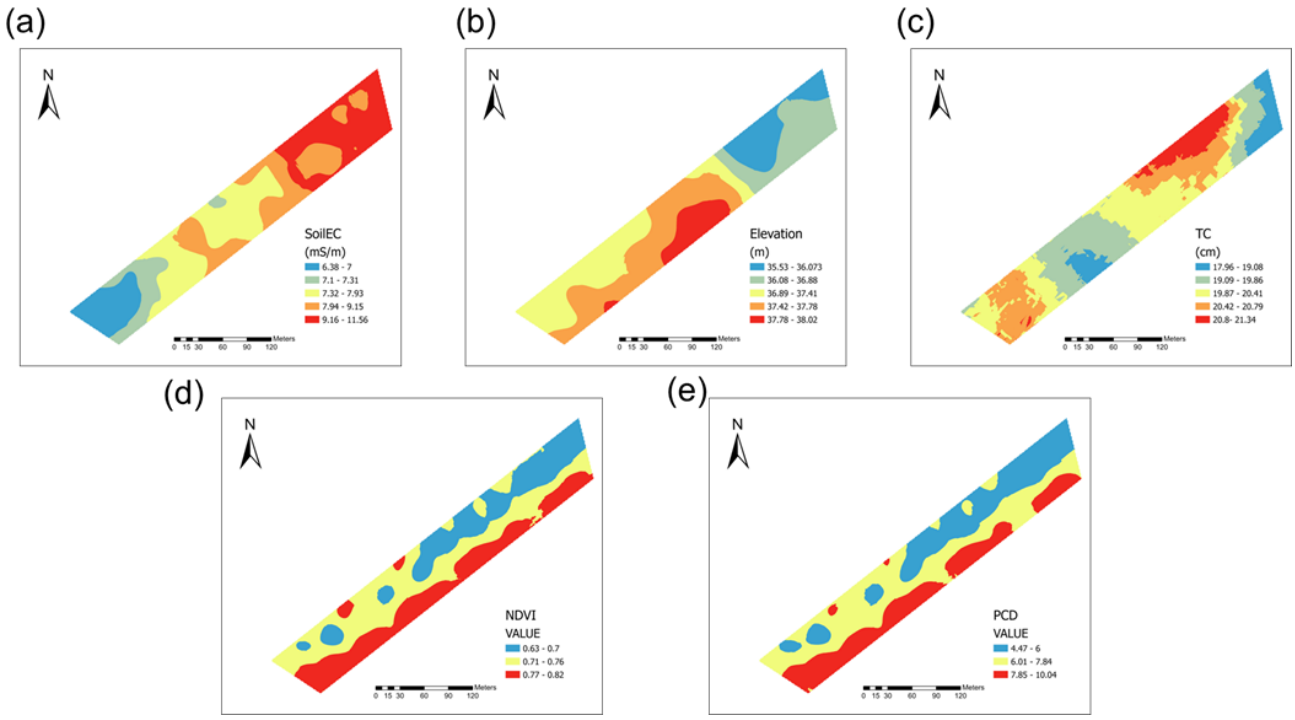

Figure 3.7: Spatial interpolation map of soil EC_a (a); elevation (b); trunk circumference (c); NDVI (d); PCD (e) in HN vineyard

Among the nonlinear methods, this study evaluated the prediction performance of the individual machine learning models (KNN, SVR), ensemble learning methods (RFR, XGBoost) and deep learning method (ANN) to estimate grape TSS. The ensemble machine learning methods including RFR and XGBoost improved the prediction performance (Table 3.5). The best prediction performance for the nonlinear methods was the RFR algorithm with an average R^2 of 0.52 and RMSE of 1.19 °Brix, respectively (Table 3.5). The result of using XGBoost showed a similar prediction performance with RFR (R^2 = 0.52, RMSE = 1.2 °Brix). RFR and XGBoost aggregated the predictions from multiple decision trees together, to significantly improve the estimation accuracy than the other machine learning models (Figure 3.8). However, the KNN model showed a significant poor prediction accuracy compared with other linear and nonlinear regression models with an average R^2 of 0.24 and RMSE of 1.2 °Brix (Table 3.5, Figure 3.8). Compared with the linear regression models, SVR significantly improved the prediction performance with an average R^2 of 0.37 and RMSE of 1.38 °Brix (Table 3.5, Figure

3.8). However, ANN showed a similar prediction performance as linear regression model with an average R^2 of 0.31 and RMSE of 1.43 °Brix (Table 3.5, Figure 3.8).

Table 3.5: The prediction performance of study machine learning models (bold represents the best prediction performance)

Method	R^2	RMSE
Lasso regression	0.3 ± 0.08	1.44 ± 0.1
Ridge regression	0.31 ± 0.08	1.43 ± 0.1
KNN	0.24 ± 0.08	1.5 ± 0.1
SVR	0.37 ± 0.08	1.38 ± 0.09
RFR	**0.52 ± 0.06**	**1.19 ± 0.07**
XGBoost	0.52 ± 0.06	1.2 ± 0.09
ANN	0.31 ± 0.11	1.43 ± 0.13

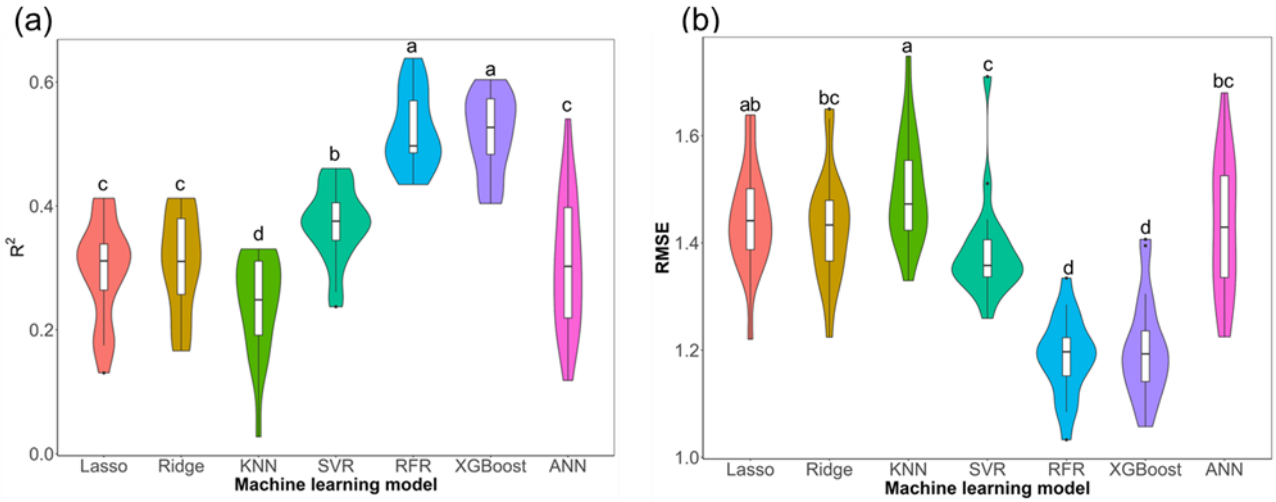

Figure 3.8: The boxplot of different machine learning model performance in R^2 (a); RMSE (b). (Different letters between any two groups represents significant difference between them, if two groups have the same letter then this indicates that they are not statistically different)

In order to reduce redundant information about the 23 VIs, OSAVI and NGBDI were used as the main input for different machine learning models with other ancillary variables. OSAVI and NGBDI consider the different scenarios of available sensors, in order to better target practical applications. Compared to regression using all 23 VIs, the result of most of machine learning models increased when using OSAVI-based model or NGBDI-based model (Table 3.6-3.7). When using OSAVI-based models, the best prediction model is RFR with an average R^2 of 0.51 and RMSE of 1.19 °Brix, respectively (Table 3.6). For NGBDI-based model, the RFR model showed best average result of predicting TSS were a R^2 of 0.54 and a RMSE of 1.16 °Brix respectively (Table 3.7). Figure 3.9 and 3.10 showed that the variation of different machine learning models' performance during the repeated 20 times different data splits strategies. The result of Waller-Duncan showed that using ensemble learning model (RFR and XGBoost) demonstrated greater capability than that of lasso, ridge, KNN, and SVR (Figure 3.9, 3.10).

Table 3.6: The prediction performance of OSAVI-based models (bold represents the best prediction performance)

Method	R^2	RMSE
Lasso regression	0.3 ± 0.08	1.44 ± 0.1
Ridge regression	0.32 ± 0.1	1.38 ± 0.1
KNN	0.4 ± 0.07	1.3 ± 0.1
SVR	0.36 ± 0.07	1.39 ± 0.1

RFR	**0.51 ± 0.07**	**1.19 ± 0.07**
XGBoost	0.5 ± 0.07	1.21 ± 0.08
ANN	0.45 ± 0.12	1.26 ± 0.14

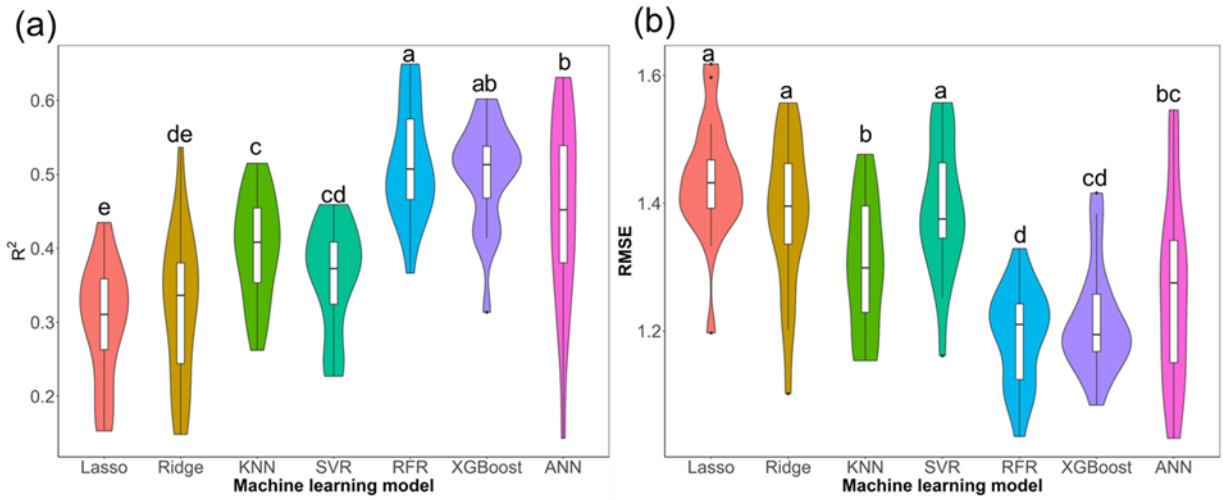

Figure 3.9: The boxplot of different OSAVI-based model performance in R^2 (a); RMSE (b). (Different letters between any two groups represents significant difference between them, if two groups have the same letter then this indicates that they are not statistically different)

Table 3.7: The prediction performance of NGBDI-based models (bold represents the best prediction performance)

Method	R^2	RMSE
Lasso regression	0.3 ± 0.08	1.43 ± 0.09
Ridge regression	0.32 ± 0.1	1.43 ± 0.09

KNN	0.4 ± 0.06	1.31 ± 0.1
SVR	0.4 ± 0.07	1.33 ± 0.12
RFR	**0.54 ± 0.07**	**1.16 ± 0.07**
XGBoost	0.52 ± 0.06	1.19 ± 0.07
ANN	0.39 ± 0.11	1.31 ± 0.11

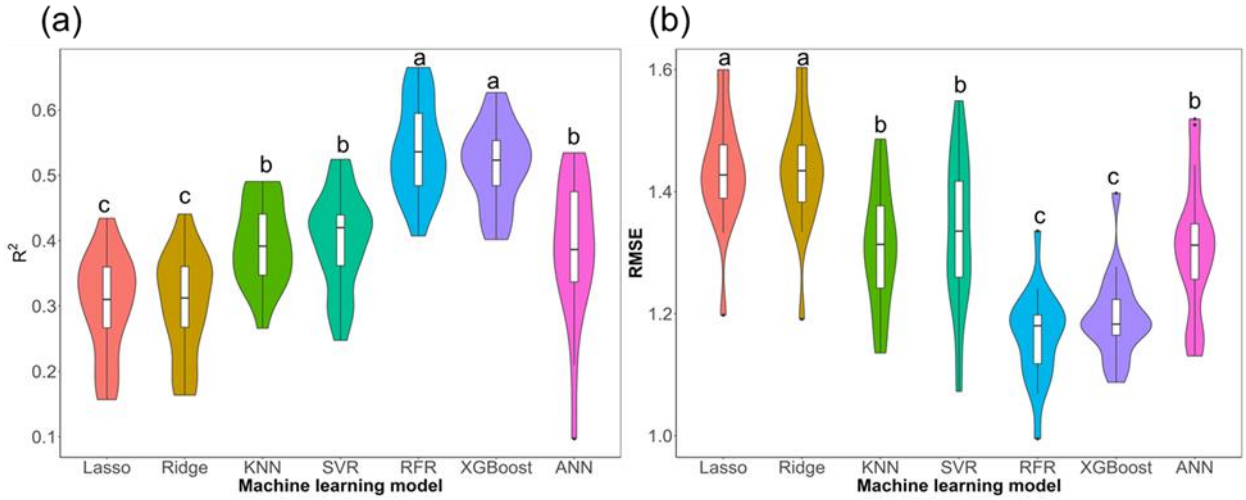

Figure 3.10: The boxplot of different NGBDI-based model performance in R^2 (a); RMSE (b). (Different letters between any two groups represents significant difference between them, if two groups have the same letter then this indicates that they are not statistically different)

3.4 Discussion

In this study, we explored the potential to use VIs, soil EC_a, elevation, slope, and TC data as input variables to predict the grape sugar content in a non-destructive way. A total of 236 samples from Pinot Noir cultivars had the TSS measured values based on destructive methods from two commercial vineyards and used as output variable in the regression model. The grape TSS was measured on five different days in the period from veraison to harvest. During the

veraison stage, the berries start to mature changing color, softening, accumulating sugar and reducing acid (Keller, 2020). From veraison, grape growers start to measure the grape quality parameters such as TSS, pH, and titratable acidity in order to determine the best harvest day. Among them, the TSS is an important parameter to assess the grape maturity as it can determine the alcohol concentration and flavor of the subsequent wine. Figure 3.4 shows that the sugar content of grapes initially decreased and then increased during the study period. However, in a previous study, the sugar concentration of grapes exhibited a strong increase starting from veraison and eventually reached a plateau during harvest stage (Trought and Bramley, 2011). One possible reason is that the sampling vines, were randomly selected during each of the five-measurement days, without repeating the selection. This differed to the sampling strategy used in the previous study. Several studies have shown that there is considerable spatial and temporal variation in grape TSS (Baluja et al., 2013; Bramley, 2005). Thus, the grapes in different geographic locations may accumulate sugar at different rates (Baluja et al., 2013). On each measurement day, the large magnitude of the interquartile range and the outlier in Figure 3.4 showed the large spatial variability in grape TSS within the vineyard blocks. Thus, it is inappropriate to take a single or average measurements, collected in the vineyard to represent the grape maturity stage.

Due to the spatial variability of grape quality at the vineyard scale, it is important to measure the grape quality across the entire vineyard. However, the traditional method relies on destructive measurement, which makes it impossible to measure each grape berry's quality. Thus many studies have explored the potential of using advanced sensing techniques to measure the grape quality with a non-destructive method (Benelli et al., 2021; Gomes et al., 2017a; Kalopesa et al., 2023; Kasimati et al., 2021). However, most of these studies used direct measurements of berries or clusters to estimate grape quality parameters (Benelli et al., 2021; Gomes et al., 2017a; Kalopesa et al., 2023; Kasimati et al., 2021). Few studies have predicted grape quality by canopy or at the leaf level (Kasimati et al., 2021). In this study we propose an alternative method to predict grape TSS based on VIs and other ancillary data. Different biological stress conditions can cause changes in canopy status such as heightened pigment

levels or canopy structure change, which will affect the quality of subsequent berries. These changes can affect the way plants interact with light of different wavelengths (Sanches et al., 2014). Based on these characteristics, many studies have used hyperspectral and multispectral images to predict the crop growth status, yield and quality (Lyu et al., 2023a). Compared with hyperspectral cameras, this study used a cheaper multispectral UAV imaging system to acquire the canopy reflectance data in the red, red edge, blue, green, and near-infrared bands. The red, red edge and near infrared bands have been reported to be related with the chlorophyll content and leaf structure. The blue and green bands have been proven to be associated with the canopy pigments change. In this study, we calculated 23 VIs based on different bands reflectance data values. Due to the grape TSS being measured from different locations without a repeat at each sampling date. The data was grouped based on measurement date and the Pearson's coefficient calculated between VIs and grape TSS in each group (Figure 3.5). A previous study calculated the Pearson's correlation coefficient between spectral indices acquired from UAV RGB imagery and grape quality parameters (García-Fernández et al., 2021). The result found significant correlations between berry weight, malic acid, alpha amino nitrogen, phenolic maturation index, total polyphenol index and spectral indices. However, the spectral indices calculated from RGB imagery show a poor correlation with TSS in their study. In this study, multispectral cameras can provide additionally spectral values in the near infrared and red-edge band. The VIs including OSAVI, RDVI, EVI, ARI, and MSAVI which are computed based on the spectral bands including near infrared and red edge, which showed a strong correlation between grape TSS during the harvest stage (Figure 3.5). Furthermore, the simple linear regression between each 23 VIs and grape TSS during harvest stage was examined. It is apparent that the prediction performance of using only RGB bands was lower than that of using the VIs which combined with near infrared and red edge bands.

The MSAVI, RDVI and OSAVI were calculated from different wavelength reflectance in the red and near infrared bands. Red and near infrared band are the most common band combination to monitor biomass and vegetation density as well as biophysical parameters (Giovos et al., 2021). PCD and NDVI calculated based on the reflectance value in red and near

infrared band were widely used in PV (Bramley et al., 2019; Bramley, 2022; Kasimati et al., 2021; Proffitt and Malcolm, 2005). For example, one study used PCD as an indicator to represent the vine vigor and explore their spatial variability within vineyard (Bramley et al., 2019). In addition, the NDVI are commonly used to identify the vine row space using high spatial resolution imagery (Hall et al., 2003). Due to the discontinuity of the vegetation surface in the vineyard, it is necessary to extract the vegetation information from the vine canopy under the high-resolution images, to reduce the influence of soil and weeds between rows. In this study a simple threshold method on NDVI and DSM-DTM, was used to identify the vine row area. However, the Pearson's correlation coefficient between NDVI and grape TSS is low in this study. A previous study showed that the NDVI obtained at veraison stage to have a strong correlation with grape TSS (Kasimati et al., 2021). One possible reason is that the UAV imagery was acquired during the flowering stage. Most studies showed that there was a strong correlation between NDVI and grape quality parameters at the late development stages of vine growing (Anastasiou et al., 2018; Sun et al., 2017). It is worth noting that, the vineyard will use the antibird net to protect the grapes after veraison. This makes it difficult to use UAV and satellites to obtain accurate canopy reflection data during this period.

Furthermore, the potential of using machine learning models to predict the grape TSS was evaluated. In addition to the VIs collected during flowering, the input variables in the machine learning includes the NDVI_proximal value measured on a handheld proximal sensor during the post veraison stage, the soil EC_a value measured from EM38-MK2, elevation and slope value obtained from LiDAR data, and TC measured in the field. The dataset was used to train and test machine learning models, evaluating the performance of linear and nonlinear regression models including lasso regression, ridge regression, KNN, SVR, RFR, XGBoost and ANN. When we used all input variables in machine learning models, the ensemble method, which included RFR and XGBoost showed similar prediction accuracy for grape TSS prediction, with the best-fitted model achieving R^2 = 0.52 and RMSE = 1.19 °Brix. These results confirm the findings of (Kasimati et al., 2021), who compared the linear and nonlinear regression models in predicting grape TSS, and they concluded that the AdaBoosting, RFR, and Extra

Trees model outperform the other machine learning models (Kasimati et al., 2021). In addition, another study showed that the XGBoost and RFR demonstrated greater capability for modeling crop yield than linear regression model and ANN (Jiang et al., 2022). Compared with (Jiang et al., 2022), this study tested different machine learning models, 20 times with different test sets, and used Waller-Duncan to analyze the differences between the performance of each model. Furthermore, this study used OSAVI or NGBDI as main input variable with other ancillary data to predict grape TSS based on different machine learning models. We choose OSAVI and NGBDI as the main input variable, as they represent the VIs can be obtained when different sensors are available. Similar results were obtained with 23 VIs used as input variables, the RFR showed the best prediction performance (Table 3.5-3.7). Therefore, the implementation of ensemble learning techniques provides potential to predict the grape TSS in a non-destructive way based on remote sensed and ancillary data. However, it should be noted that berry quality was affected by the environmental conditions during harvest stage (e.g. radiation, temperature). The further study should continue to explore using different input variables to predict grape TSS. In addition, the berry samples were only 3 berries per vine which cannot represent the whole vine grapes' TSS value, the few samples size may influence the prediction performance when using different machine learning models. Thus, further study should consider improve the berry sampling numbers.

3.5 Conclusions

This study investigates the possibility of using the combination of remote sensed and ancillary data to predict the wine grape quality parameter used as a proxy for TSS. The Pearson's correlation coefficient showed that the VIs obtained from the UAV during the flowering stage have a strong correlation with the grape TSS during the harvest stage. The input variables which include VIs obtained from UAV, NDVI_proximal and soil EC_a obtained from proximal sensors, elevation and slope obtained from a LiDAR camera, as well as TC obtained directly in the vineyard was used to build regression models to estimate grape TSS in a non-destructive way. This study evaluates the prediction performance of seven machine

learning techniques: ridge regression, lasso regression, KNN, SVR, RFR, XGBoost and ANN. The result shows that ensemble learning models (RFR, and XGBoost) outperform other regression models when predicting grape TSS. This study develops an alternative approach to predict the grape TSS by different predictor values through various advanced techniques. For grape growers, the approach developed in this study could help them assess the whole vineyard grape TSS in a non-destructive way, in order to make the harvest decision.

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Hongyi Lyu
Name and title of main supervisor:	Dr Miles Grafton, Senior Lecture
In which chapter is the manuscript/published work?	Chapter 4

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]
The candidate used field spectroscopy to predict grape sugar content in three commercial vineyards during two harvesting seasons.
The supervisory panel assisted with laboratory assistance and other technical support, editing and writing support.

Please select one of the following three options:

○ The manuscript/published work is published or in press
Please provide the full reference of the research output:

◉ The manuscript is currently under review for publication
Please provide the name of the journal:
Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2024). Non-destructive and on-site estimation of grape total soluble solids by field spectroscopy and stack ensemble learning. European Journal of Agronomy (with editor)

○ It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	Hongyi Lyu 2024. 06. 29	Main supervisor's signature:	Miles Grafton Digitally signed by: Miles Grafton DN: CN = Miles Grafton email = m.grafton@massey.ac.bz C = NZ O = Massey University OU = SAE Date: 2024.06.29 13:57:19 +12'00'

This form should be placed at the beginning of each relevant thesis chapter.

[1] Refer to the Massey University Publishing and Authorship guidelines (OneMassey for staff, Stream for students) and/ or Contributor Roles Taxonomy (CRediT) guidelines for guidance.

4 Non-destructive and On-site Estimation of Grape Total Soluble Solids by Field Spectroscopy and Stack Ensemble Learning

Using remote sensing imagery and ancillary variables showed a moderate prediction performance to predict grape TSS value during the pre-harvest stage. One possible explanation is that the strong spatial variability of grape TSS value within a vine. The small number of berries taken from each vine does not represent the quality and ripeness of the entire vine. After veraison, leaf growth stops, and berries start to soften and accumulate sugar. During this stage, grape growers commonly use nets to shield berries from birds, complicating the use of remote sensing techniques for monitoring berry quality status. Proximal sensors such as VNIR-SWIR spectroscopy provide an alternative to assess the berry quality. To date, existing methodologies primarily rely on controlled laboratory conditions and spectroscopic analysis in the VNIR region. Consequently, there exists a gap in the application of using proximal sensing techniques to predict grape TSS based on VNIR-SWIR field spectra data. Additionally, no conclusive research has been done to determine whether it is better to use spectra data in VNIR or SWIR, or VNIR-SWIR, to predict grape quality. In order to address some of these shortcomings, this chapter aims to explore the comparative effectiveness of VNIR, SWIR, and VNIR-SWIR spectral data in predicting grape TSS under unstable environmental conditions.

This chapter based on:

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2024). Non-destructive and on-site estimation of grape total soluble solids by field spectroscopy and stack ensemble learning. *European Journal of Agronomy* (with editor)

Abstract

Accurately estimating the total soluble solids (TSS) of berries with a non-destructive method is crucial for wine grape growers if wine quality improvements are to be made. At present, the methods employed with the best statistical results are implemented under stable lab conditions, using spectroscopic analysis in the visible-near infrared (VNIR) region. This study explores using field spectroscopy to estimate the TSS of berries directly in the vineyard. A portable visible-near infrared-shortwave infrared (VNIR-SWIR) spectroradiometer measured the reflectance data of grape berries in the 350-2500 nm spectral region. A large in-field multi-season spectra database (n = 1830) over two years (2023-2024) from three 'Pinot Noir' commercial vineyards were selected to develop spectral-region specific (VNIR, SWIR or VNIR-SWIR) machine learning models. Different machine learning modeling pipelines were built using data collected from 2023 and validated using data from 2024 to predict grape TSS based on in-field spectral databases. Subsequently, the performance of using stack ensemble learning (ES) to predict grape TSS was evaluated and compared with three commonly used methods: K-nearest neighbors (KNN), random forest regression (RFR), and support vector regression (SVR). The result on the independent test set showed that, the ES model based on first derivative spectra data, in the VNIR-SWIR region provided the highest prediction accuracy for grape TSS value, with a coefficient of determinations (R^2) of 0.803, root mean square error (RMSE) of 1.213 °Brix, and a ratio of performance to deviation (RPD) of 2.086, with a Lin's concordance correlation coefficient (CCC) of 0.89. This study demonstrated that grape growers could implement an ES model to assess the grape TSS rapidly and non-destructively from field spectroscopy data.

Keywords: machine learning; grape berries; field spectroscopy; spectroradiometer; total soluble solids

4.1 Introduction

Grapevine (*Vitis* spp.) is one of the key berry crops in New Zealand, due to its commercial use in wine production. Accurately estimating the quality of berries is crucial for industrialized wine-growing systems, as the quality of wine is strongly influenced by the grape quality at the time of picking (Chen et al., 2015). Grape quality is traditionally monitored by TSS in New Zealand. TSS in grape pulp can be measured using lab-based destructive techniques such as utilizing a refractometer to measure their °Brix value. The traditional measurement method only provides an average representation of the TSS in a specific area of the vineyard, while previous studies have shown that TSS varies spatially in a single block (Baluja et al., 2013). Winemakers regard the uniformity of fruit berries as a key aspect of wine quality. When uniform berries are utilized, wine can be produced with the consistent flavor and aroma. This is more profitable than selling low quality wines because a relatively small portion of the input fruit does not have the desired attributes (Bramley, 2022). To obtain the spatial-temporal variation of TSS in fruit throughout the vineyard, repeated measurements must be taken at various locations and times during ripening. This requires many destructive measurements over each vineyard. However, this is not desirable for commercial vineyards as it is labor intensive and expensive. Therefore, in the commercial wine industry, there is a significant demand for non-destructive methods that can offer fast and dependable results for assessing grape TSS (Dambergs et al., 2015).

Many studies have explored the possibility of using spectroscopic analysis in assessing various grape quality parameters or varieties non-destructively (Larraín et al., 2008; González-Caballero et al., 2010, 2012; Urraca et al., 2016; dos Santos Costa et al., 2019; Benelli et al., 2021; Kalopesa et al., 2023). Most of these studies utilized spectroscopic measurement to predict TSS in wine grape berries under controlled laboratory conditions and obtained high prediction accuracy (dos Santos Costa et al., 2019). However, the application of spectroscopy in vineyard conditions faces challenges due to unpredictable environmental factors such as light, temperature, and sample location. Larraín et al., (2008) and Guidetti et al., (2010) demonstrated the potential of using field spectroscopy for monitoring grape quality parameters

directly in the vineyard. However, the spectral range of reflectance data used in these studies is in the VNIR region (400-1000 nm). The utilization of spectral reflectance data, particularly in the SWIR region (1000-2500 nm), has gained considerable interest in grape production, as indicated by González-Caballero et al., (2012). Their study revealed that the SWIR spectra, particularly in the wavelength range of 1600-2400 nm, can be an effective tool in classifying the ripening stage of wine grapes. In addition, Urraca et al., (2016) achieved moderate prediction accuracy, with RMSE of 1.68 °Brix for TSS using a spectra range of 1600-2400 nm. In studies conducted in the VNIR-SWIR region, Kalopesa et al., (2023) achieved good prediction accuracy, with a R^2 of 0.87 (RMSE of 1.76 °Brix) for TSS using a convolutional neural networks model. To the best of our knowledge, there is currently no definitive research indicating whether the spectra data in VNIR or SWIR, or a combination of both (VNIR-SWIR), is preferable for predicting TSS. Costs present a crucial consideration for industrial applications, with VNIR spectroscopy being generally cheaper when compared to SWIR. Therefore, it is crucial to evaluate the effect of different spectral regions on predicting TSS.

Furthermore, machine learning techniques have been employed to leverage the rich information present in spectral reflectance data, enabling the identification of relationships between reflectance and fruit quality parameters (Kalopesa et al., 2023). Nevertheless, the high dimensionality and multicollinearity issues in spectral reflectance data can lead to overfitting of machine learning models, reducing their accuracy and generalizability usefulness. It is essential to perform effective data pre-processing and feature selection, to mitigate these issues and enhance the performance, of machine learning algorithms when working with hyperspectral data. Dimensionality reduction methods are commonly employed in studies to reduce storage space, computation time, and remove redundant features in high-dimensional data, in order to mitigate multicollinearity issues (Silva and Melo-Pinto, 2023, 2021). The dimensionality reduction methods can be broadly classified into two categories: linear techniques, such as Principal Component Analysis (PCA), and non-linear techniques, such as t-Distributed Stochastic Neighbor Embedding, Kernel PCA and autoencoders. Silva and Melo-Pinto, (2021) applied various dimensionality reduction techniques in a machine learning model to predict

grape sugar content and found PCA outperformed other non-linear techniques. Their approach depended on single predictive models, SVR and are subject to overfitting with limited training data. Over the past decade, various ensemble techniques that use a collection of base learners to improve predictive performance were proposed. Common ensemble strategies include bagging, boosting, and stacking. Bagging generates base learners in parallel by creating training subsets through bootstrap sampling. Boosting trains a sequence of base models by capitalizing on their dependences (Breiman, 1996). Stacking, unlike bagging and boosting, typically employs heterogeneous learners and leverages their differences to improve accuracy (Wolpert, 1992). By ensuring diversity among the base learners, complementary information is provided, thus the risk of incorrect results is reduced. To the best of our knowledge, the application of a stacking ensemble strategy in grape quality prediction has not been observed, though it has been shown in successful applications in predicting crop yield in various studies on other crops (Ji et al., 2023).

This study utilizes field spectroscopy data to identify the TSS value of wine grapes in a non-destructive way. The specific objectives of this research are:

- To develop a non-destructive method to estimate wine grape TSS based on field spectroscopy.
- To compare the performance of stack ensemble learning model with individual machine learning model.
- To evaluate the effect of different spectral regions (VNIR, SWIR, and VNIR-SWIR) on model performance.
- To quantify the effects of different data pre-processing methods on model estimations.
- To validate model transferability in different seasons.

4.2 Methodology

4.2.1 Study Site

The study was conducted during the period between veraison and harvest seasons of 2023 and 2024 (early February to late March) at three commercial vineyards: Hua Nui (HN) (both years 2023-24), Pencarrow (PN) (in year 2023), and Clouston (CN) (in year 2023) (Figure 4.1). The vines were > 15 years old 'Pinot Noir' and trained with vertical shoot positioning. The soils in the vineyards are mostly clay and silt loams, which are known to have moderate soil water holding capacity. The vineyard managers were wholly responsible for cultivation practices uniformly across the sample areas of each vineyard.

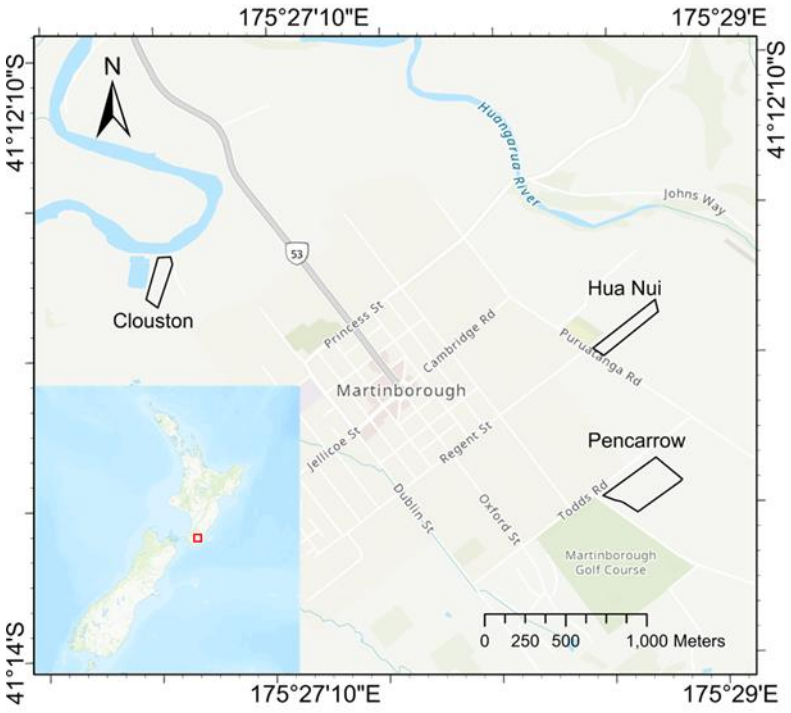

Figure 4.1: Location of study vineyards (The red square represents the study site)

In 2023 and 2024, a total of 1843 grape berries were collected to estimate TSS. In 2023, seven field visits were carried out in three vineyards (HN, PN, and CN). Whereas in 2024, four field visits were carried out in HN vineyard (Table 4.1). On each measurement date, several

healthy vines were selected across the vineyards. Each sampling vine provided berries selected from the top, middle and bottom position of a single bunch (Figure 4.2a).

Table 4.1: Berry sampling information collected from each vineyard (where n = number of samples)

Years	Locations	Sampling dates	N	Vine spacing (m)	Row spacing (m)
2023	HN	20/02/2023	75	1.7	2.2
		26/02/2023	108	1.7	2.2
		02/03/2023	40	1.7	2.2
		06/03/2023	90	1.7	2.2
		14/03/2023	42	1.7	2.2
	PN	20/02/2023	26	1.8	2.2
		26/02/2023	78	1.8	2.2
		02/03/2023	109	1.8	2.2
		06/03/2023	96	1.8	2.2
		14/03/2023	131	1.8	2.2
		24/03/2023	33	1.8	2.2
	CN	24/03/2023	153	1.6	2.2
		28/03/2023	128	1.6	2.2
2024	HN	06/02/2024	41	1.7	2.2

		23/02/2024	231	1.7	2.2
		03/03/2024	231	1.7	2.2
		09/03/2024	231	1.7	2.2

4.2.2 Acquisition of Spectral Reflectance Data and Pre-Processing

In-situ spectral acquisitions were performed using the ASD FieldSpec 4 Hi-Res NG Spectroradiometer (Malvern Panalytical Ltd., Malvern, UK), which is capable of measuring the VNIR-SWIR spectrum (350-2500 nm). The spectroradiometer employed a contact probe for controlled illumination (Figure 4.2b). During the measurement, a white panel ceramic referencing tile was utilized for calibration, and as a reference spectrum, after each set of three measurements on a given vine. The ratio of the sample's optical energy to that of the reference panel was used to calculate reflectance values. Reflectance spectra data of each individual berry were obtained when they grew in grape clusters in reflectance mode by averaging five measurements. The spectral data were subsequently interpolated to achieve a spectral resolution of 1 nm, resulting in a dataset comprising 2151 bands ranging from 350 to 2500 nm. To eliminate signal noise, reflectance data outside the spectral regions below 450 nm and above 2200 nm were excluded. The processed data were then exported as ASCII text files for

each spectral measurement using ViewSpec Pro 6.2 software (Analytical Spectral Devices, Inc., Boulder, CO, USA).

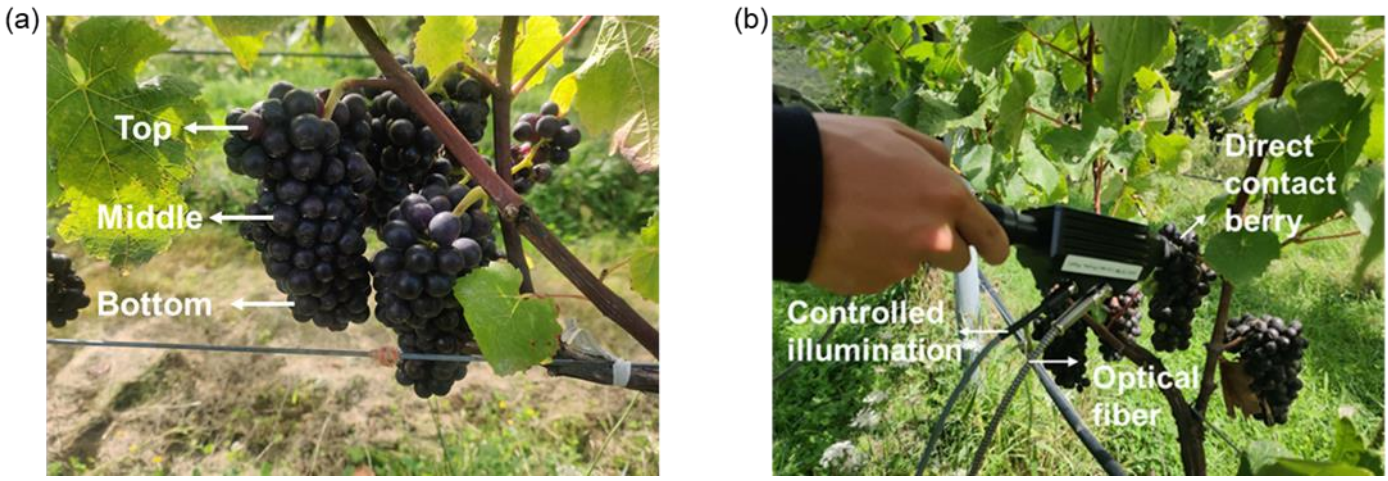

Figure 4.2: The location of sampled berries in a grape cluster (a), Collection of the grape berry spectra using the ASD FieldSpec 4 Hi-Res NG Spectroradiometer (b)

To enhance the subsequent modeling phase and eliminate physical phenomena present in the spectra, the raw reflectance data (RW) were subjected to various spectral transformations. Specifically, the first derivative (1D), Savitzky–Golay filtering with polynomial = 3, window size = 11, and derivative = 0 (SG), and absorbance (Abs = Log (1/RW)) were applied. These transformations aimed to mitigate factors such as baseline shifts and slope variations, which could affect the modeling process (Rinnan et al., 2009). Derivative pre-processing procedures were performed using the “prospectr” package in R statistical software (R Core Team, version 4.2.2). A derivative gap of 3 was employed during the computation of the derivatives (Wei et al., 2021; Lyu et al., 2023a).

4.2.3 TSS Content Measurement

Following the acquisition of spectral reflectance data, the measurement of TSS in grape berries was carried out on the same grape berries which were used for acquiring spectral reflectance data. TSS, expressed in °Brix, was measured by using a portable digital refractometer (PAL-ALPHA Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan). The

refractometer had a measuring range of 0-85 °Brix with an accuracy of 0.20 °Brix. Prior to each measurement, the refractometer prism underwent a thorough rinse with deionized water, followed by drying using absorbent cleaning paper. Calibration was performed by adjusting the instrument to read zero degrees of °Brix using a drop of deionized water at around 25 °C. For analysis, each grape berry selected was carefully cut and squeezed to extract the juice, which was then placed onto the transparent prism surface of the refractometer.

4.2.4 Dimensionality Reduction Method

The dimensionality reduction method improves the prediction performance of machine learning models for spectral reflectance data by removing redundant features in a dataset (Silva and Melo-Pinto, 2021). In this study, PCA was chosen for the dimensionality reduction method for the three feature groups (RW, 1D, SG, and Abs). The PCA has been widely utilized in dimensionality reduction of spectral reflectance data and proved to outperform other methods (Silva et al., 2018; Silva and Melo-Pinto, 2021).

PCA enabled the transformation of high-dimensional data into a lower-dimensional space, while preserving as much of the original variation as possible. The transformation results in the creation of “principal components (PCs)” which were linear combinations of the original features. In general cases, the optimal number of PCs to retain for analysis is determined by the proportion of variance explained (selecting the number of PCs which accounted for the majority of the variance in the population). However, there is no one best method for determining how many PCs should be retained for analysis. Thus, in this study we set the PCs as a hyperparameter in different machine learning modeling pipelines. Consequently, the model was tested for each dataset using a range of one to 50 PCs, and the ideal number of PCs was selected based on cross-validation (Silva et al., 2018). We choose 50 PCs as the upper limit since the 50 PCs explain 99.99% of the variability in the original spectra data. Additionally, retaining a higher number of PCs for analysis leads to an increase in computational costs.

4.2.5 Data-Driven Modeling

Before the application of the machine learning model, anomalous samples detection was performed based on 1.5 times interquartile range (IQR) rule and PCA. The outlier in the spectra data was detected by the Mahalanobis distance of the first and second principal components in the PCA. The outlier in the °Brix data was detected by the limits of 1.5 IQR. The anomalous samples were excluded when the limits of 1.5 IQR and PCA were crossed simultaneously. In this case 1830 samples were selected for further analysis. To estimate TSS from spectral reflectance data (RW, 1D, SG, and Abs), the following machine learning algorithms were applied: KNN, RFR and SVR. KNN regression is an instance-based learning method that predicts the numerical target based on distance functions (e.g., Euclidean and Mahalaobis). KNN regression as a non-parametric method, predicting the response by calculating the average of the responses from the KNN in the training set. Delgadillo-Duran et al., (2022) successfully used KNN to predict soil properties from spectral data. Additionally, RFR and SVR are widely used in spectroscopic analysis to predict grape quality parameters (Tsakiridis et al., 2023). RFR is an ensemble learning algorithm that generates a series of regression trees by using different bagged samples. The final predicted value of RFR is obtained by averaging the predictions generated by all the individual regression trees in the ensemble. SVR is an extension of the support vector machines specifically designed for regression problems. It finds the best hyperplane to maximize the margin between the predicted values and the observed values, making it useful for predicting continuous numerical values.

In addition, an ES model was used in order to improve the prediction performance. The ES model is a common hierarchical ensemble learning strategy, that typically comprises two levels: level 1, consisting of basic learners; and level 2, a meta-learner. The basic learners encompass multiple machine learning algorithms, and the outputs from these basic models were collected and combined to generate meta-features. The basic learners in this study included the KNN, RFR and SVR. The highest accuracy performance of the basic learner was selected as the meta-learner for the ensemble model (Figure 4.3). For processing all machine learning models, the mlr3verse package in R software version 4.2.2 was used. Appendix 2 shows hyperparameter settings of different machine learning models

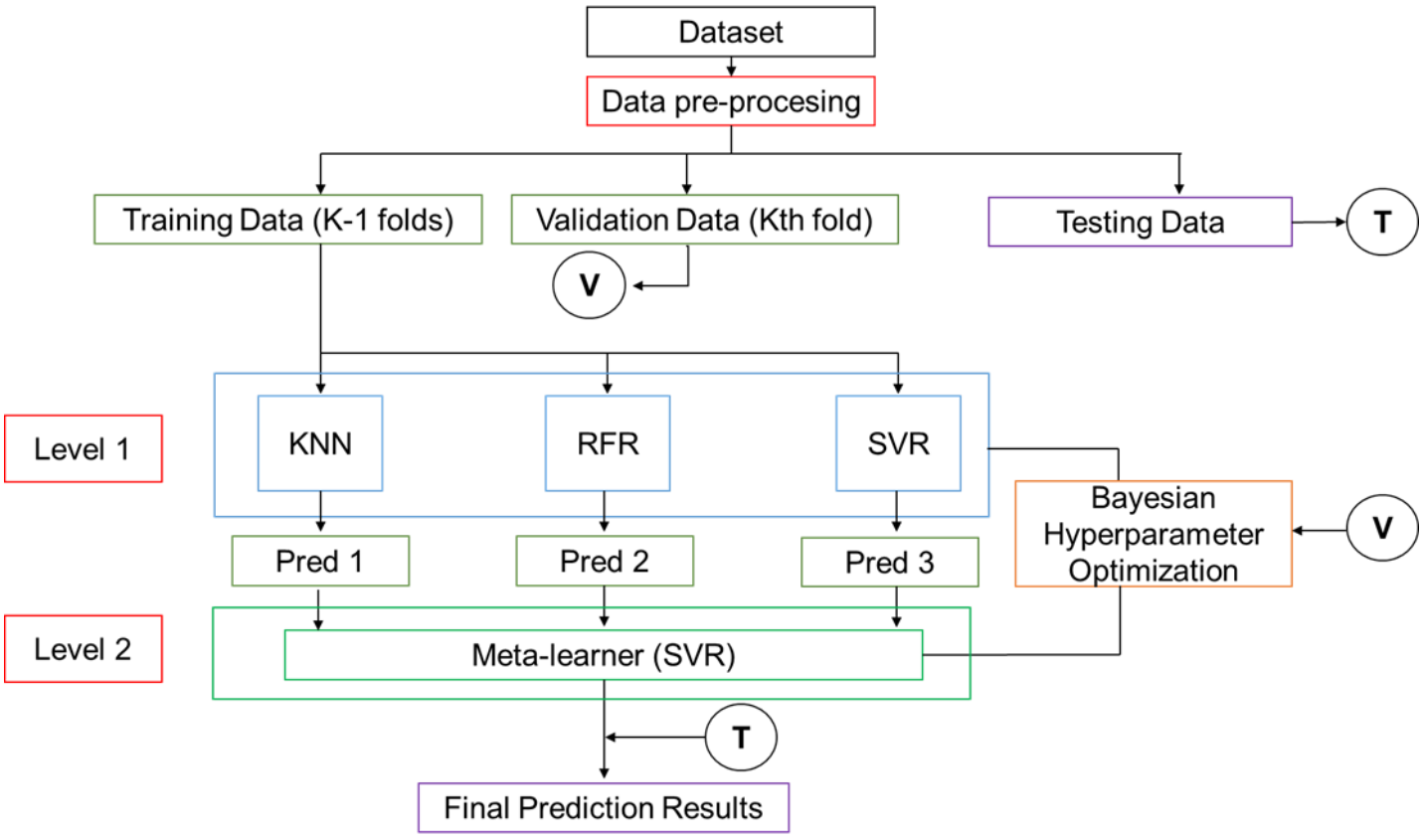

Figure 4.3: The workflow of the ES model for TSS estimation

4.2.6 Model Performance Evaluation

In this study, the train-test split method was implemented according to the year of data collection. Data collected in 2023 (n = 1099) were designated for model training, while data

collected in 2024 (*n* = 731) were reserved for model performance evaluation. This approach employs an independent dataset for validation, reducing the risk of overfitting, and evaluating the generalization performance of the model. Additionally, a 10-fold cross validation strategy was used to tune the hyperparameters in each training set of all the tested machine learning models. The combination of hyperparameters contributing to the models with the lowest RMSE values, were considered as optimized. In order to evaluate the model performance, the values of RMSE, ratio of performance to deviation (RPD), R^2, and Lin's concordance correlation coefficient (CCC) were calculated by applying the trained models with optimized hyperparameters to the test set. The equations of the indices are shown in Equations 4.1 – 4.4.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad \text{Equation 4-1}$$

$$\mathrm{RPD} = \frac{SD(y)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}} \qquad \text{Equation 4-2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad \text{Equation 4-3}$$

$$\mathrm{CCC} = \frac{2\rho\sigma_{y_i}\sigma_{\hat{y}_i}}{\sigma_{y_i}^2 - \sigma_{\hat{y}_i}^2 + (\mu_{y_i} - \mu_{\hat{y}_i})^2} \qquad \text{Equation 4-4}$$

where n is the number of samples used to fit the model, y_i is the ground truth value of the ith sample, $\bar{y}$ is the mean response value, $\hat{y}_i$ is the model estimated value of the ith sample.

4.3 Result

4.3.1 Statistical Analysis of Measured TSS and Spectral Reflectance Data

The period from veraison to harvest is recognized as the most crucial phase for monitoring the quality of grape berries (González-Caballero et al., 2012). After outlier detection, throughout this period, a total of 1830 berries were collected, and the summary statistics of the training and testing datasets on berry TSS values is presented in Table 4.2. In this study, a wide range of TSS value were found. Berry TSS ranged from 8.2 to 30.9 °Brix, indicating that many grapes reached maturity (Silva et al., 2018). The wide range of the dataset made it easier to build a robust calibration model. Table 4.2 showed high TSS variability (SD = 2.53) for the samples

collected from HN vineyard in 2024 compared to the samples of the other vineyards; whilst low TSS variability (SD = 1.59) for the samples from CN vineyard.

Table 4.2: Descriptive statistics of training and testing datasets on total soluble solids (where n = number of samples, min = minimum, max = maximum, SD = standard deviation)

Year	Datasets	Vineyard	n	min	median	max	SD
2023	Train	HN	352	8.3	17	23.4	2.02
		PN	466	8.2	16.9	21.2	2.11
		CN	281	16.1	20.3	24.8	1.59
2024	Test	HN	731	9.7	20	30.9	2.53

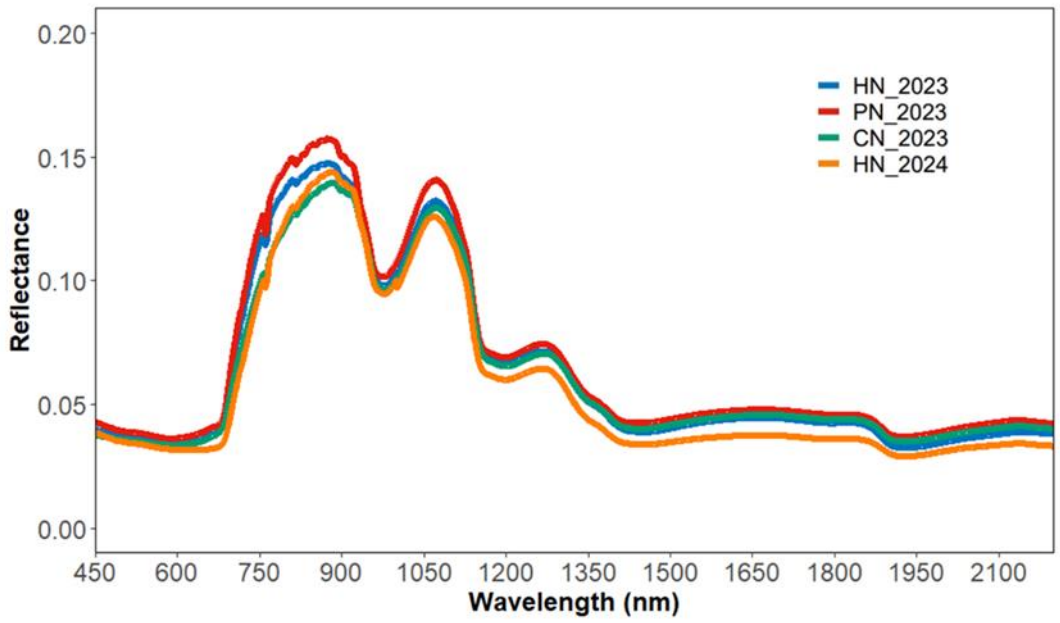

Figure 4.4: The mean spectral values of 'Pinot Noir' grape samples from different vineyards in different seasons

Figure 4.4 showed the average spectra of the grape samples collected from different vineyards in different seasons, covering the spectral range of 450–2200 nm. The small reflectance between 500 and 700 nm was related to pigments such as anthocyanins, chlorophyll, and carotenoids present in fruit (Ncama et al., 2017). The absorption at around 970 nm was related to carbohydrate and water. González-Caballero et al., (2010) found that the 940-980 nm region as the most relevant for TSS measurement. In addition, sugar-related absorption bands were also found at around 1200 nm (Williams and Norris, 1987). However, most of the short-wave infrared (SWIR, 1000–2500 nm) regions appeared to provide limited information in terms of absorption bands. The weak absorption at around 1450 nm and 1940 nm was related to water in fruit (Fairuz Omar, 2013).

PCA was conducted to reduce the redundant information in the spectra data. After PCA, the original spectra data were transformed into principal components. According to the PCA analysis the first three principal components accounted for 71.4%, 21.4% and 3.9% of the variance, respectively (Figure 4.5a, b). Figure 4.5c illustrated the loadings plot for the first five principal components. There are a large number of peaks throughout the spectrum from 700-1000 nm, that are usually associated with sugar content.

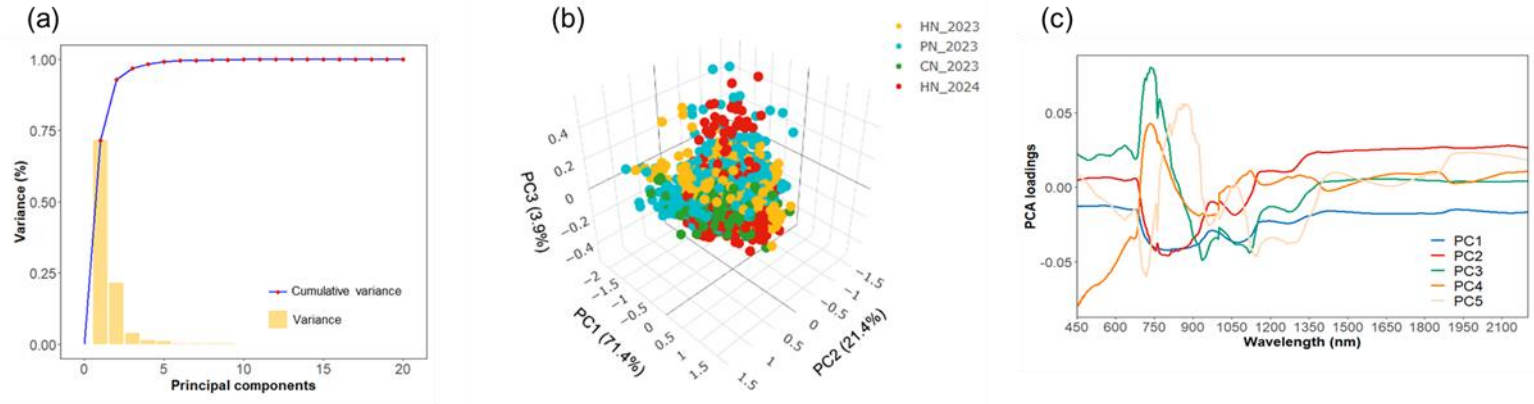

Figure 4.5: The result of principal component analysis of spectra data. Cumulative variance of the principal components (a), 3D-scored images of the first three principal components (b), Loadings plot for the first five principal components (c)

4.3.2 Prediction Accuracy Based on Different Machine Learning Models

All the machine learning models (KNN, RFR, SVR and ES) were trained using different pre-processing methods and tested on independent datasets. The best results of different machine learning models based on different spectral pre-processing and dimensionality reduction methods are shown in Table 4.3. The model performance criteria showed that the ES model (R^2 = 0.803, RMSE = 1.213 °Brix, CCC = 0.89 and RPD = 2.086) outperformed KNN (R^2 = 0.634, RMSE = 1.53 °Brix, CCC = 0.783 and RPD = 1.654), RF (R^2 = 0.688, RMSE = 1.413 °Brix, CCC = 0.814 and RPD = 1.791), and SVM (R^2 = 0.765, RMSE = 1.225 °Brix, CCC = 0.882 and RPD = 2.065) models on the independent test set (Table 4.3). The scatter plot of predicted and measured TSS value on the independent test set for each best performance machine learning models are shown in Figure 4.6.

Table 4.3: Model performance criteria for predicting TSS from field spectroscopy data on the independent test set includes coefficients of determination (R^2), root mean square error (RMSE), ratio of performance to deviation (RPD), and Lin's concordance correlation coefficient (CCC)

Model	Spectral region	Pre-processing	R^2	RMSE	CCC	RPD
KNN	VNIR	1D	0.634	1.53	0.783	1.654
RF	VNIR	Abs	0.688	1.413	0.814	1.791
SVM	VNIR-SWIR	1D	0.765	1.225	0.882	2.065
ES	VNIR-SWIR	1D	0.803	1.213	0.89	2.086

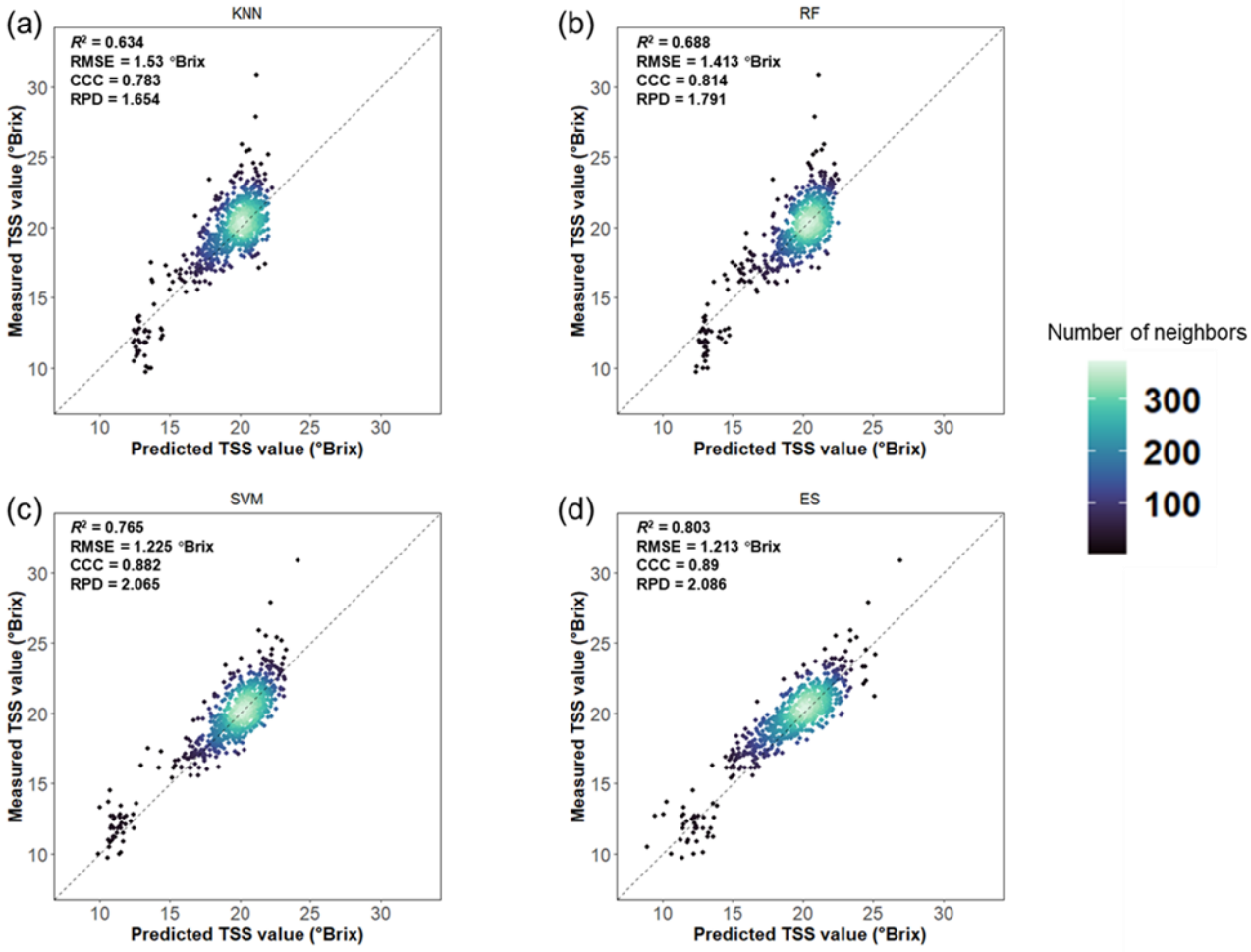

Figure 4.6: The 1:1 line relationship between predicted and measured TSS values for the independent test set by KNN based on 1D spectra data in the VNIR region (a), RF based on absorbance spectra data in the VNIR region (b), SVM based on 1D spectra data in the VNIR-SWIR region (c), ES based on 1D spectra data in the VNIR-SWIR region (d)

The result of using different spectra region data for predicting TSS values shows the TSS estimates apparently favor the VNIR region for KNN and RF models as compared to the SWIR or VNIR-SWIR spectral regions (Figure 4.7). However, SVM and ES models showed higher accuracy in the VNIR-SWIR spectral region when predicting TSS (Figure 4.7). The result of RMSE from different machine learning models and pre-processing methods showed a consistent result where the prediction performance using SWIR region alone are lower than those using VNIR and VNIR-SWIR (Figure 4.7).

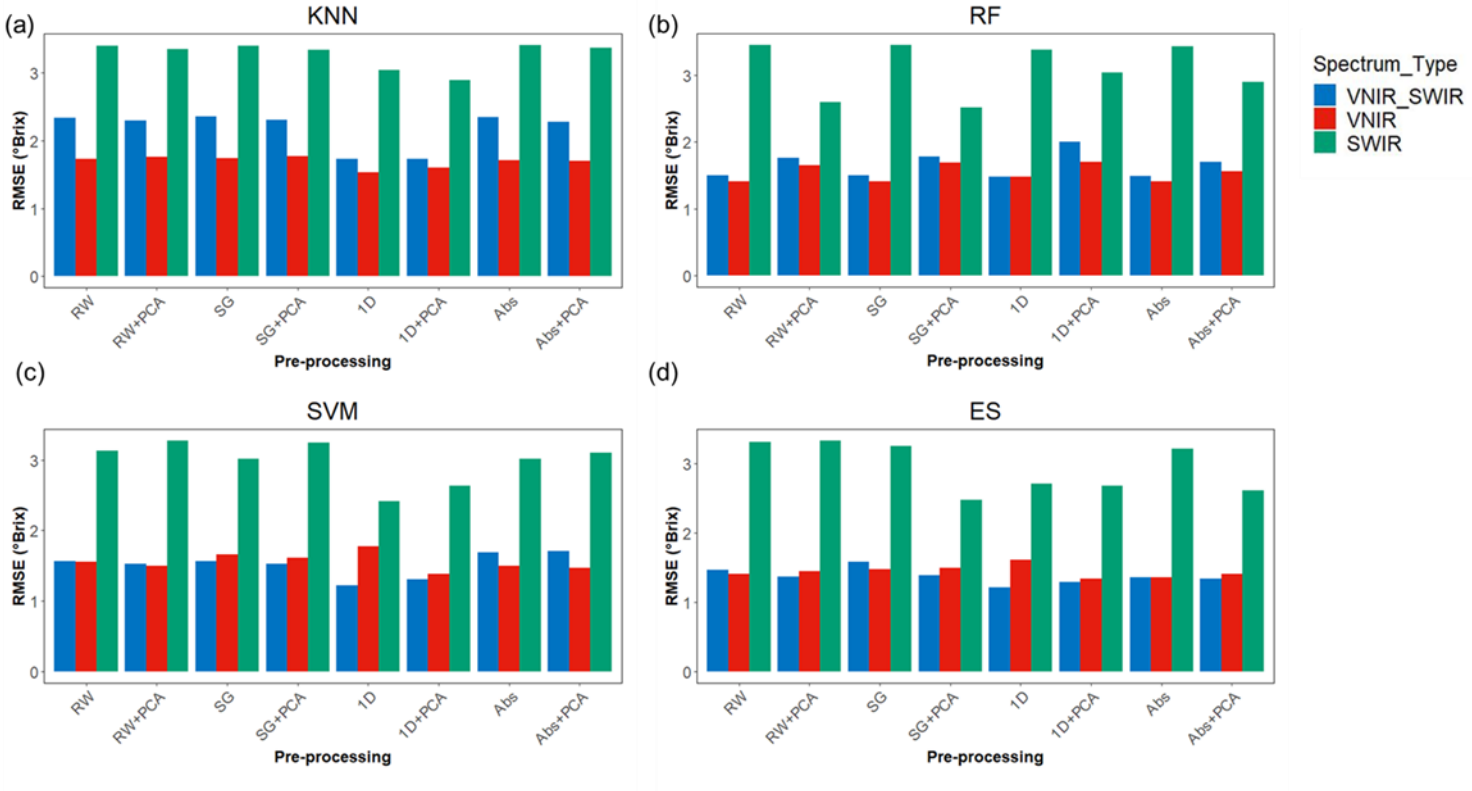

Figure 4.7: The TSS prediction accuracy (RMSE) of the four machine learning models (KNN, RF, SVM and ES) using the grape berry spectral features in the VNIR, SWIR and VNIR-SWIR spectral regions for independent test set

4.4 Discussion

It was labor intensive to manually acquire each berry quality measurement from each vine. Traditional destructive measurement only measures a few berries in a single block to represent the whole vineyard's berry quality. The present study explores a novel non-destructive approach to estimate the TSS of wine grapes when berries are grown in grape clusters through a portable VNIR-SWIR spectroradiometer (350-2500 nm). Different machine learning models and pre-processing methods were applied to predict TSS from in-field spectral data. To assess the generalization of the machine learning models, different seasonal data sets were used as an independent test set.

4.4.1 The Stack Ensemble Learning Model Performance on TSS Estimation

Several studies have utilized machine learning models in combination with spectra data to predict grape berry quality (Fernandes et al., 2015; Silva et al., 2018). For example, SVR a commonly non-linear machine learning models has been applied to predict grape berry TSS and pH (Silva et al., 2018). In addition, Gomes et al., (2017) have demonstrated the application of neural networks as an individual prediction model for predicting grape TSS. It is worth noting that these studies relied on an individual basic learner for constructing predictive models, related to grape berry quality. In order to improve the prediction accuracy, this study used a stacking-based ensemble learning (ES) approach, which combines three individual learners to construct the TSS prediction. KNN, RFR and SVR were utilized for constructing ES, and SVR was selected as the meta-learner for this ensemble approach. The result of using the ES model with different pre-processing methods and spectral region data outperformed other individual machine learning models in independent test sets (Table 4.3, Figure 4.7). Compared with individual machine learning models, the ES model can reduce bias and variance in the final prediction by capturing intricate patterns that individual models cannot capture alone. Several successful examples of use of the ES approach have also been observed in crop yield prediction (Ji et al., 2023). In general, this study evaluates 96 machine learning pipelines for predicting grape TSS value. The result shows the best prediction performance was achieved by applying ES model to 1D data in the VNIR-SWIR spectral region (R^2 = 0.803, RMSE = 1.213 °Brix, CCC = 0.89 and RPD = 2.086), which was better than the findings of Gutiérrez et al., (2019) where the RMSE was 1.274 °Brix. This study demonstrates the success of using the ES approach to predict the berry quality.

4.4.2 Influence of Spectral Region on Model Performance

Silva & Melo-Pinto, (2021) suggested that the VNIR spectral region is optimal for predicting grape TSS because of its cheaper acquisition Most studies measured the grape berry reflectance in the VNIR (400-1000 nm) spectral region alone to predict grape TSS values (Guidetti et al., 2010; Silva et al., 2018; Tsakiridis et al., 2023; Benelli et al., 2021). However,

dos Santos Costa et al., (2019) and Kalopesa et al., (2023) measured reflectance data in the VNIR-SWIR spectral region. The prediction results from dos Santos Costa et al., (2019) and Kalopesa et al., (2023) were slightly better than that of other studies using VNIR spectral regions to predict grape quality. This study explores the influence of different spectral regions (VNIR, SWIR, VNIR-SWIR) on model performance. Different machine learning models show the best predictions in different spectral regions (Table 4.3, Figure 4.7). SVM and ES model in this study have shown slightly higher performance within the VNIR-SWIR spectral region. However, the KNN and RF model showed lower accuracy within the VNIR-SWIR spectral region to that of the VNIR region. For the best performance model (ES model), a wider range of spectral regions are required, rather than using spectral data only in VNIR region, which improves the model predictions. However, compared with VNIR, the use of VNIR-SWIR spectroscopy for grape TSS prediction does not significantly improve the prediction performance, so in practical applications, wine grape growers are currently recommended to use VNIR equipment rather than VNIR-SWIR because of its present lower cost and good prediction ability.

4.4.3 Influence of Pre-processing Method on Model Performance

In order to solve the high dimensionality and multicollinearity issues within the spectral reflectance data, this study used different pre-processing methods to reduce the dimensionality of the data to an appropriate level. The PCA approach was widely employed as one of the most commonly used dimensionality reduction methods. It offers an acceptable performance while requiring moderate computational resource (Silva and Melo-Pinto, 2021). The efficiency of using transformed variables from the PCA method for predicting TSS was evaluated. The ES model performance in the independent test set showed that using the PCA method to pre-process spectra data can improve the model performance among most model pipelines in the VNIR-SWIR spectral region (Figure 4.7), which was consistent with previous studies (Silva et al., 2018; Silva and Melo-Pinto, 2021). Despite PCA not enhancing prediction accuracy across all models, its application offers a substantial reduction in model training time. By condensing high-dimensional spectral data into fewer components, PCA lowers computational demands,

which can be particularly beneficial in scenarios with limited processing resources. Future research should explore using additional dimensionality reduction techniques such as independent component analysis (ICA), t-distributed stochastic neighbour embedding (t-SNE), and autoencoders in model pipeline.

Previous studies showed that different spectral pre-processing can improve the prediction accuracy when predicting grape berry TSS, grape leaf nutrient status, and grape water status (Kalopesa et al., 2023; Lyu et al., 2023a; Wei et al., 2021). In this study, when using KNN, RFR, SVR and ES algorithms, it was observed 1D transformations improve the modeling performance compared to models trained with raw reflectance data. However, other preprocessing methods do not improve the performance of the model in most cases (Figure 4.7). One possible reason for the poor performance of transformation variables is that since the spectral pre-processing method used in this study did not consider the scatter correction method. In future studies, more combinations of spectral pre-processing and dimensionality reduction method should be selected to investigate the impact of pre-processing methods on predicting berry quality.

4.4.4 Model Generalization

Most studies restricted their study area to one vineyard or one season (dos Santos Costa et al., 2019; Gutiérrez et al., 2019; Benelli et al., 2021; Kalopesa et al., 2023). The test and train data set in these studies often come from the same environmental conditions, thus high model performance can be expected. However, these models are less likely to extend to an independent data. Only a few studies have explored the transferability of models, seeking to create a model from one vineyard and utilize it on different vineyards or harvest season (Urraca et al., 2016; Silva et al., 2018). Independent validation stands out as the most reliable approach for evaluating a model's generalization ability. Thus, this study used an independent dataset which was collected in a different harvest season to evaluate four machine learning models' generalization ability. The result showed a satisfactory range of the prediction performance and model generalization ability of four machine learning models (Table 4.3). Comparing the results

in the previous studies using in-lab spectra data from different vintages in the testing phase, we found that Silva et al., (2018) obtained a R^2 of 0.837, and a RMSE of 2.443 °Brix, for the determination of TSS, when using different varieties and vintages as test sets. Even though spectral data used in this study were collected from unstable environmental conditions (Figure 4.2), the model performance (R^2 = 0.803, RMSE = 1.213 °Brix, CCC = 0.89 and RPD = 2.086) is similar to Silva et al., (2018) when using the ES model. This indicates robustness and reliability of the ES mode in processing field spectra data. This study presents a solution for using field spectroscopy to replace conventional techniques for predicting grape TSS value. The results confirm the suitability of predicting 'Pinot Noir' TSS values. Non-destructive and rapid estimate of grape TSS values aids vineyard managers to determine harvest strategies such as selective harvesting and mapping spatial variability of grape TSS. Note that the independent dataset was collected in a different harvest season, the vineyard used for independent testing is one of the three vineyards included in the training dataset. Future studies should test the current ES model across vineyards in a wider range of growing environments and grape varieties to enhance its robustness and applicability .

4.5 Conclusion

Estimating the grape berry TSS value during harvest is important for vineyard management to ensure better berry quality. This study has explored a large database collected in different vineyards and harvest seasons to establish a simple, relatively less computational costly, and non-destructive method for predicting grape berry TSS directly in-field using different machine learning modeling pipelines. The results revealed that the ES model based on 1D spectra data in the VNIR-SWIR region was suitable to estimate 'Pinot Noir' TSS values (R^2 = 0.803, RMSE = 1.213 °Brix, CCC = 0.89 and RPD = 2.086). In order to better apply this method to the real-world industry, further research should consider validating this model to more grape varieties. The developed methodology in this study has demonstrated the possibility for a rapid and non-destructive measurement of TSS in wine grapes based on field spectroscopy.

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Hongyi Lyu
Name and title of main supervisor:	Dr Miles Grafton, Senior Lecture
In which chapter is the manuscript/published work?	Chapter 5

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]

The candidate used hyperspectral imaging system to predict grape quality parameters in lab conditions based on different regression and classification models.
The supervisory panel assisted with a hyperspectral camera and other technical support, editing and writing support.

Please select one of the following three options:

- (●) The manuscript/published work is published or in press

  Please provide the full reference of the research output:
  Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2024). Hyperspectral Imaging Spectroscopy for Non-Destructive Determination of Grape Berry Total Soluble Solids and Titratable Acidity. Remote Sensing, 16 (10), 1655. https://doi.org/10.3390/rs16101655

- ( ) The manuscript is currently under review for publication

  Please provide the name of the journal:

- ( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	Hongyi Lyu 2024. 06. 29	Main supervisor's signature:	Miles Grafton Digitally signed by: Miles Grafton DN: CN = Miles Grafton email = m.grafton@massey.ac.bz C = NZ O = Massey University OU = SAE Date: 2024.06.29 13:59:18 + 12'00'

This form should be placed at the beginning of each relevant thesis chapter.

[1] Refer to the Massey University Publishing and Authorship guidelines (OneMassey for staff, Stream for students) and/ or Contributor Roles Taxonomy (CRediT) guidelines for guidance.

5 Hyperspectral Imaging Spectroscopy for Non-Destructive Determination of Grape Berry Total Soluble Solids and Titratable Acidity

Compared with spectroscopy techniques, hyperspectral imaging system can provide spatial resolution of traits in grape berries. After harvesting, wine grape enters into a grape de-stemmer to remove the berries from the rachis. At this stage, a hyperspectral imaging system can be combined with a grape de-stemmer to achieve selective harvesting. This chapter is dedicated to evaluating using spectra data acquired from a hyperspectral imaging system to predict and classify grape quality parameters during the post-harvest stage. The ability of various spectral preprocessing techniques and machine learning models to predict and classify grape total soluble solids and titratable acidity was investigated, and the contributing spectral features that had the most influence were identified.

This chapter based on:

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2024). Hyperspectral Imaging Spectroscopy for Non-Destructive Determination of Grape Berry Total Soluble Solids and Titratable Acidity. *Remote Sensing*, *16*(10), 1655. https://doi.org/10.3390/rs16101655

Abstract

Wine grape quality heavily influences the price received for the product. Hyperspectral imaging has the potential to provide a non-destructive technique for predicting various oenological parameters. The study aims to explore the feasibility of applying hyperspectral imaging to measure total soluble solids (TSS) and titratable acidity (TA) in wine grape berries. A normalization different spectral indices (NDSI) spectral pre-processing method was built and compared it with the conventional pre-processing method: multiplicative scatter correction and Savitzky-Golay smoothing (MSC+SG). Different machine learning models were built to examine the performance of pre-processing methods. The result showed the NDSI pre-processing method showed higher performance than the MSC+SG pre-processing method in different classification models, with the best model correctly classifying 93.8% of TSS, and 84.4% for TA. In addition, TSS can be predicted with moderate performance using support vector regression (SVR) and MSC+SG pre-processing with root mean squared error (RMSE) of 0.523 °Brix, and coefficient of determinations (R^2) of 0.622, the TA can be predicted with moderate performance using SVR and NDSI pre-processing (RMSE = 0.19%, R^2 = 0.525). This study demonstrated that hyperspectral imaging data and NDSI pre-processing has the potential to be a method for grading wine grapes for producing quality wines.

Keywords: wine grape quality, non-destructive methods, machine learning, hyperspectral imaging

5.1 Introduction

The quality of wine is strongly related to the compounds of the grape during harvest. Traditionally, grape maturity is monitored by measuring grapes' TSS and TA values. The TSS of grape berries directly affects the potential alcohol concentration, fermentable sugar concentration and flavor profile of the subsequent wine (Gomes et al., 2017a). TA are the concentration of organic acids present in the berries, such as tartaric acid, malic acid and citric acid, which are responsible for the overall sour taste of the subsequent wine (Lee et al., 2013). These two oenological parameters determine the optimal timing for grape harvest and, consequently, the quality of the resulting wine. Currently, the measurement of grape TSS and TA is destructive, by means of handheld refractometers or traditional chemical analysis in laboratory settings. Due to differences of position and orientation of the berries within each berry cluster, the quality of the berries vary spatially at the cluster level (Baluja et al., 2013). Most growers adopt a uniform harvesting approach, collecting grapes of varying quality together and processing them into wine. However, this production method may diminish winery profitability, as the inclusion of lower quality berries during the production process can diminish the overall wine quality and subsequently lower its market value (Bramley et al., 2003). To obtain the spatial variation of grape quality would require a large number of destructive measurements using the traditional method. The use of destructive measurements of TSS and TA on a large number of grapes is not feasible for wine industry production, as normally viticulturists sample 200 berries per block to represent the whole block grapes’ quality. Consequently, using a non-invasive and a rapid method such as hyperspectral imaging (HSI) could provide a method to measure the TSS and TA on large numbers of grapes.

HSI is a non-destructive technique that combines conventional imaging with spectroscopy to capture and analyze the spatial and spectral information of objects (M. Zhao et al., 2023). The HSI is commonly used in reflectance, transmittance and interactance modes. Among these modes, the reflectance mode is the most frequently used for analyzing grape berries (Li et al., 2019). In reflectance mode, the light source directly illuminates the object's surface, and the

detector records information regarding the intensity of the reflected light from objects, which varies according to the wavelength. Several studies have investigated using laboratory-based HSI technology to determine the following oenological parameters: TSS, TA, pH and anthocyanins (Chen et al., 2015; Gomes et al., 2017a; Silva et al., 2018). One study used a laboratory based HSI system in reflectance mode to predict grape TSS in Portugal (Gomes et al., 2017a). Through the development of prediction models based on PLSR and neural networks, prediction performance (R^2) up to 0.95 was achieved. The laboratory based HSI system, combined with a SVR model was employed to predict anthocyanin concentration, pH index and TSS in whole grape berries. Determination coefficient (R^2) of 0.89 (RMSE = 35.6 mg. L^{-1}), 0.81 (RMSE = 0.25), and 0.90 (RMSE = 3.19 °Brix) were obtained for the estimation of anthocyanin concentration, pH index and TSS respectively. These two studies highlight the potential of using HSI to predict the oenological parameters in a rapid and non-invasive way. However, it is worth noting that the HSI systems commonly used in existing literatures primarily operate within the visible (VIS) and near infrared (NIR) region (400 - 1000 nm) (Gomes et al., 2017a; Silva et al., 2018). Several studies explored the reflectance responses of grape berries in the regions of short-wave infrared (SWIR, 1400 - 3000 nm) and indicated the potential for predicting oenological parameters in NIR and SWIR region (Ferrer-Gallego et al., 2011; González-Caballero et al., 2012; Chen et al., 2015; Urraca et al., 2016). One study used a handheld spectroradiometer with spectral ranging from 1600 to 2400 nm to predict the grape TSS within a lab setting (Urraca et al., 2016). The resultant prediction performance yielded R^2 value of 0.91 and RMSE value of 1.42 °Brix.

Most previous studies used regression models to predict grape TSS or TA, while the main objective of determining the grape quality is distinguishing berries into two groups (ripe and unripe berry). The classification algorithms are important to distinguish the quality of grapes. For example, one study explored spectroradiometer (350-2500 nm) combined with various supervised classification methods to distinguish the maturation stages of wine grapes and obtained 93.15% of accuracy (dos Santos Costa et al., 2019). In addition, several studies have reported using spectroradiometer to classify grape quality based on TSS or TA (Dambergs et

al., 2003; Guidetti et al., 2010). Compared with spectroradiometer, HSI system can provide more spatial information and is also easy to install on a berry sorting system for industrial application. Therefore, it is important to explore the potential of using the combination of HSI and classification algorithms to distinguish grape quality.

Machine learning has been widely applied in analyzing hyperspectral data, as hyperspectral imaging can provide a large amount of information with high complexity. However, many machine learning models are called “black box” as the decision-making process of these models is often deemed uninterpretable (Lyu et al., 2023a; Wei et al., 2021). Many interpretation methods have been developed aimed to identify the features that contribute most to the model's predictive performance. In grape studies, this can be useful to better understand the relevance of the specific wavelength for predicting different oenological parameters (Lyu et al., 2023a).

This study aims to develop and evaluate non-destructive and rapid prediction of two important compositional parameters named TSS and TA using NIR-SWIR hyperspectral imaging. The specific objectives were: (1) to develop regression models to predict grape TSS and TA using the combination of HSI and machine learning algorithms; (2) to develop classification models to distinguish grape quality; (3) to identify the spectral signatures which contribute most to the model’s performance.

5.2 Materials and Methods

5.2.1 Sample Acquisition

Samples of Pinot Noir grapes were collected three times during the pre-harvest stages in 2022 at the vineyard in Martinborough (41°13'8.73"S, 175°27'36.23"E), New Zealand. The vineyard features a cool, moist, and temperate climate (Lyu et al., 2023b). Random selection was employed to choose a total of 112 grapevines during the study period (Table 5.1). From each grapevine, three grape bunches were randomly selected (Figure 5.1). Subsequently, three berries were obtained from the top, middle and bottom within each bunch cluster and placed in labeled plastic bags immediately. Each sampled grapevine provided 9 berries to meet the

minimum standards for subsequent TA and TSS measurements. The collected samples were immediately sent to the laboratory for hyperspectral spectrum analysis and chemical determination of TSS and TA values. All samples were stored at 4°C until beginning the hyperspectral imaging analysis.

Table 5.1: The number of samples during each sampling date

Date	Number of samples
07th March	32
15th March	57
27th March	33

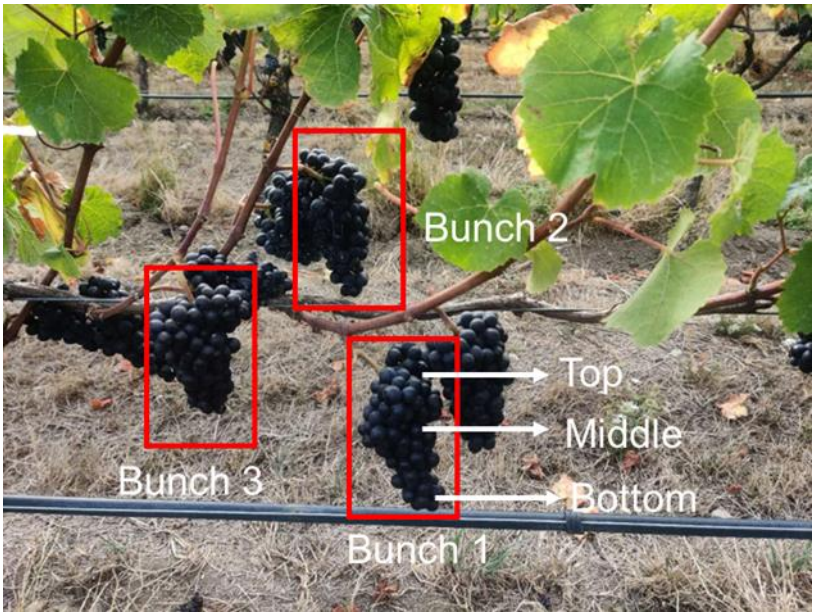

Figure 5.1: Example of the sampled grapevine

5.2.2 Hyperspectral Image Acquisition and Analysis

The Specim SWIR HSI system (Specim, Oulu, Finland) operating in reflectance mode was used to acquire images containing NIR to SWIR spectra ranging from 1000 to 2500 nm, comprising 288 spectral bands (Figure 5.2). The Specim SWIR is a pushbroom instrument which can capture images consisting of 384 spatial pixels. Two 50-watt tungsten halogen lamps were positioned at a 45-degree angle on both sides of the camera to provide illumination primarily focused on the grape samples for spatial information acquisition during scanning (Figure 5.2). The camera was situated 300 mm away from the objective base, and the imaging process was conducted using Specim Lumo software within a dark room at room temperature. Each image was acquired at a rapid rate of 29 frames per second, with an exposure time of 2 ms. The grape berries were placed on a computer-controlled conveyor belt (with the black light absorbing material), which ensured a consistent movement at a constant speed of 15 mm/s during the measurement process. A total of 112 images, each containing 9 fresh berries, were obtained using this HSI system (resulted in 112 samples).

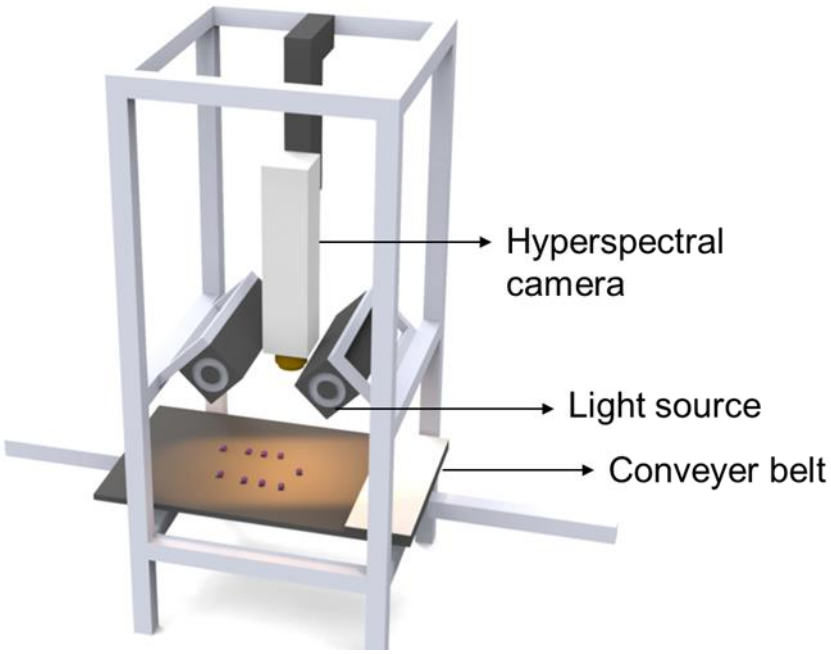

Figure 5.2: The Specim SWIR hyperspectral imaging system

To obtain surface reflectance from the raw hyperspectral images, two calibration steps were carried out using dark current and white reflectance. The dark current calibration involved closing the camera lens with a non-reflective opaque black cap to capture the dark current values. For the white reference calibration, a white panel coated with barium sulphate, known for its near-perfect surface reflectance, was used. The reflectance value is computed as:

$$R_\lambda = \frac{R_{raw\,(\lambda)} - R_{dark\,(\lambda)}}{R_{white\,(\lambda)} - R_{dark\,(\lambda)}} \qquad \text{Equation 5-1}$$

where $R_{raw\,(\lambda)}$ represents the intensity of light reflected from objective at λ wavelength, $R_{dark\,(\lambda)}$ and $R_{white\,(\lambda)}$ represent the dark current and white reflectance at λ wavelength respectively.

The hyperspectral images were processed by the Environment for Visualizing Images (ENVI) V 5.6 software (Research Systems Inc., Boulder, Co, USA). A region of interest (ROI) was defined in the middle portion of each berry and subsequently, the average spectrum of these ROI was calculated to be used as the predictor variable for the prediction model (Figure 5.3). The average value of nine grapes in each image was calculated to represent one sample. The ellipsoidal shape of the fruit resulted in the presence of shadows around the edges of each fruit, where the lighting was insufficient to fully illuminate the area. As a result, these shadowed regions were excluded from the selection of ROI.

The resulting hyperspectral image exhibited various random noise which generate from the sensor or quantization process. The noise significantly degraded the quality of spectral data and the overall prediction performance (Pullanagari and Li, 2021). Before building the estimation model, multiplicative scatter correction and the Savitzky-Golay smoothing (MSC+SG) spectral pre-processing methods, were applied to the raw reflectance data, to decrease the influence of the random noise and surface scattering. The mean of the data set was selected as reference spectrum. Additionally, normalization different spectral index (NDSI), was also used to process raw reflectance data. The NDSI was calculated as:

$$NDSI = \frac{R_x - R_y}{R_x + R_y}$$ Equation 5-2

where R_x and R_y represent the reflectance at wavelength x and y, respectively.

The Pearson's coefficient correlation between NDSI and TSS as well as TA was calculated. The Pearson's coefficient correlation ranges from +1 to -1. The closer ± 1, the stronger correlation between NDSI, TSS and TA. NDSI with coefficients higher than 0.5 selected for building the further regression and classification model.

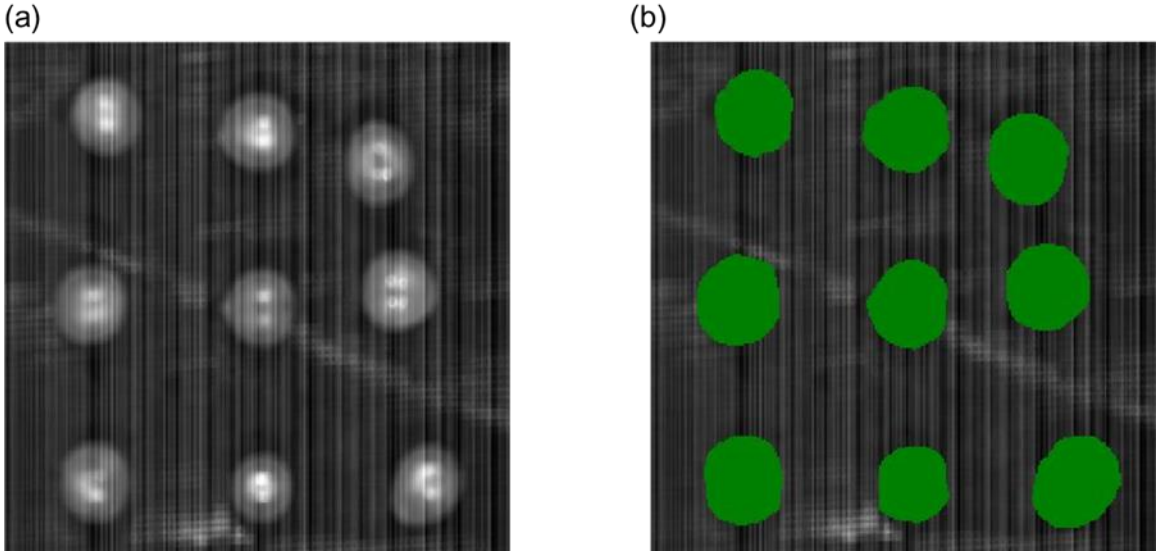

Figure 5.3: The sampled grape berries image in band 1177 nm (a); region of interest grape berries (b)

5.2.3 Oenological Parameters Measurements

After acquiring each hyperspectral image, the oenological parameters of grape berries were measured in the Massey Agri-food Digital Lab, Massey University, Palmerston North. Firstly, nine grape berries were squeezed together into juice. Then, we applied 0.2ml of the collected juice onto a portable digital refractometer (PAL-BX/ACID F5 Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan) to measure the TSS value (°Brix). The refractometer is capable of measuring 0-60% °Brix (0.1% resolution) and 0.1-8.8% (0.01% resolution) acid. After

measuring the TSS, using reverse osmosis (RO) water to rinse the refractometer prism, then dried by absorbent cleaning paper. The TA value was measured by digital refractometer. We poured 1.00g of grape juice into a beaker and diluted with RO water at a ratio of 1:50. Then we placed the 0.3ml diluted sample onto the transparent prism surface of the refractometer and recorded the acidity (%) value.

5.2.4 Data Analysis

Before the application of the prediction model, anomalous samples and spectral outlier detection was performed based on Interquartile Range (IQR) and principal component analysis (PCA). The outlier was detected by the Mahalanobis distance of the first and second principal components in PCA. In this case 107 samples were selected for building the prediction model.

5.2.4.1 Development of Regression Model

Three popular multivariate regression models partial least squares regression (PLSR), SVR and random forest regression (RFR) which have been successfully applied in HSI studies, were used to build the quantitative model for the TSS and TA content in wine grapes. PLSR projects the predictor variables onto a set of latent variables which are linear combinations of the original predictor variables. These latent variables capture the maximum covariance between predictor variables (spectral data) and outcome variable (TSS and TA). SVR is a regression technique based on support vector machines. It aims to find a hyperplane in a high-dimensional space to fit the data through kernel functions. The Gaussian radial basis function (RBF) was chosen as the kernel function in this study. RFR is an ensemble algorithm creating decision trees from diverse bagged samples, with the final prediction being the average of all individual decision trees' predictions in the ensemble.

The total samples were stratified and split into training and testing sets using a 70/30 ratio based on the distribution of each variable to evaluate the different model performance. To determine the optimal combination of hyperparameters and reduce overfitting, a grid search and 10-fold cross-validation were performed in the training set, with RMSE serving as the evaluation metric. Additionally, we used the variable importance projection (VIP) or permutation

feature importance (PFI) to calculate the contribution of each spectral wavelength in the best performance regression models among the PLSR, SVR and RFR models. The testing set was used to evaluate the performance based on the root mean squared error (RMSE), coefficient of determinations (R^2), ratio prediction to deviation (RPD) and Lin's concordance correlation coefficient (CCC). The RMSE, RPD, R^2, and CCC of the test set were calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{Equation 5-3}$$

$$RPD = \frac{SD(y)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}} \quad \text{Equation 5-4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{Equation 5-5}$$

$$CCC = \frac{2\rho\sigma_{y_i}\sigma_{\hat{y}_i}}{\sigma_{\hat{y}_i}^2 - \sigma_{\hat{y}_i}^2 + (\mu_{y_i} - \mu_{\hat{y}_i})^2} \quad \text{Equation 5-6}$$

where n is the number of samples used to fit the model, y_i is the ground truth value of the ith sample, $\hat{y}_i$ is the model estimated value of the ith sample, $\bar{y}$ is the mean response value, SD is the standard deviation, ρ is the R^2, σ is the corresponding variance, μ is the mean value.

5.2.4.2 Development of Classification Model

Distinguishing the grapes into different quality groups helps winemakers process them into wines of different qualities thus increasing the profitability of wineries. Grape hyperspectral reflectance data were divided into two groups (low or high quality) based on the TSS and TA to develop the classification model. Three popular classification models latent dirichlet allocation (LDA), support vector machine (SVM), and random forest (RF) were performed. The LDA method was implemented using linear distance definition for the computation of distance between samples to distinguish different classes. SVM and RF were the SVR and RFR used for classification analysis. The samples were divided into two groups (high and low quality) for each parameter and separated by threshold limiting values: 19 °Brix for TSS and 1.5% for TA. The same strategy of splitting the dataset into training and testing sets when developing regression models was used. In addition, a grid search and 10-fold cross-validation were performed in the training set, with accuracy serving as the evaluation metric to determine the

best hyperparameters when training the model. The testing set was used to evaluate the performance based on the accuracy (Acc), recall, false positives (FP) and classification error Rate (Ce) which calculated from the confusion matrices of each model. The calculation of Acc, recall, and Ce was defined bellow:

$$Acc = \frac{TP+TN}{TP+FN+FP+TN} \quad \text{Equation 5-7}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{Equation 5-8}$$

$$Ce = \frac{FP+FN}{TP+FN+FP+TN} \quad \text{Equation 5-9}$$

where TP, FN, FP, and TN are the true positive, false negative, false positive, true negative.

5.3 Result

5.3.1 Statistical Analysis

The descriptive statistics of measured TSS and TA obtained by the destructive method is shown in Table 5.2. Table 5.2 shows the changes of TSS and TA within a month during harvest time. The TSS value increased from 16.9 to 21.2 °Brix during the study period. At the same time, the TA decreased from 2.26 to 0.92 %.

Figure 5.4 shows the processed spectra of each sample in the spectral range 1026 - 2027 nm. The spectra for the samples showed a similar pattern. The reflectance between 1150 – 1200 nm and 1350 – 1400 nm was due to the overtone of C-H and O-H (Chen et al., 2015). Additionally, sugars such as glucose and fructose present in grapes can exhibit C-H stretching vibrations that contribute to the absorption peak at around 1200 nm (Williams and Norris, 1987). The strong reflectance peak at around 1100 nm was due to the combination bands of O-H in water (Iqbal et al., 2013). The absorption peak at around 1190, 1450 and 1940 nm were also related to the large amount of water in the berry (Fairuz Omar, 2013).

Table 5.2: Descriptive statistics of oenological parameters during study period

Parameter	Date	N	Minimum	Maximum	Mean	SD
	07th March	21	17.3	19	18.21	0.58
°Brix	15th March	53	16.9	19.9	18.63	0.67
	27th March	33	18.8	21.2	19.59	0.49
	07th March	21	1.28	2.26	1.72	0.33
TA	15th March	53	1.17	1.93	1.59	0.17
	27th March	33	0.92	1.51	1.23	0.14

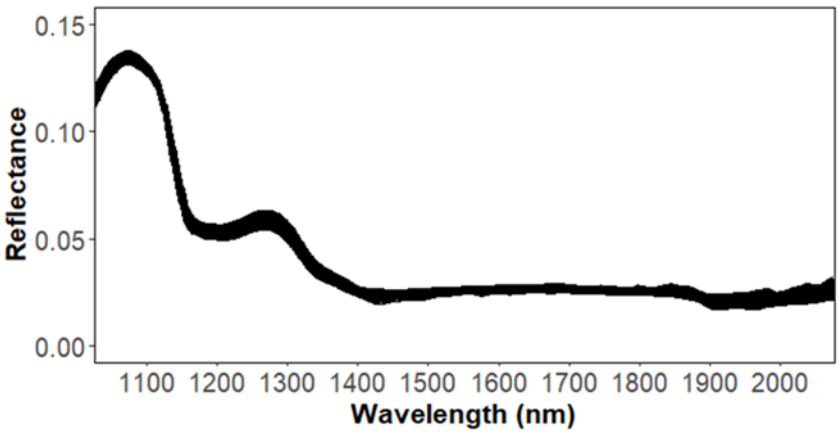

Figure 5.4: Reflectance spectrum of Pinot Noir variety

5.3.2 TSS and TA Estimation Based on Hyperspectral Narrowband NDSI

To minimize the signal noise and scattering effect, hyperspectral reflectance has been transformed into narrowband spectral indices, for example spectral ratio (SR), difference

spectral indices (DSI) and normalized difference spectral indices (NDSI) (Wang et al., 2022). Among them, NDSI has been widely used, so we explored the correlation between TSS, TA and NDSI, as shown in Figure 5.5. NDSI was calculated for each possible wavelength combinations. The highest correlation variable between TSS and NDSI was (R_{1289}-R_{1177})/(R_{1289}+R_{1177}), with a correlation coefficient of 0.62. In addition, (R_{1294}-R_{1177})/(R_{1294}+R_{1177}), (R_{1076}-R_{1060})/(R_{1076}+R_{1060}), and (R_{1300}-R_{1177})/(R_{1300}+R_{1177}), had a high correlation with TSS. The most relevant NDSI with TA was (R_{1411}-R_{1434})/(R_{1411}+R_{1434}), with a correlation coefficient of 0.6. The correlation plots showed the combination of NDSI between 1500 – 1800 nm had a higher correlation with TA than that of TSS (Figure 5.5).

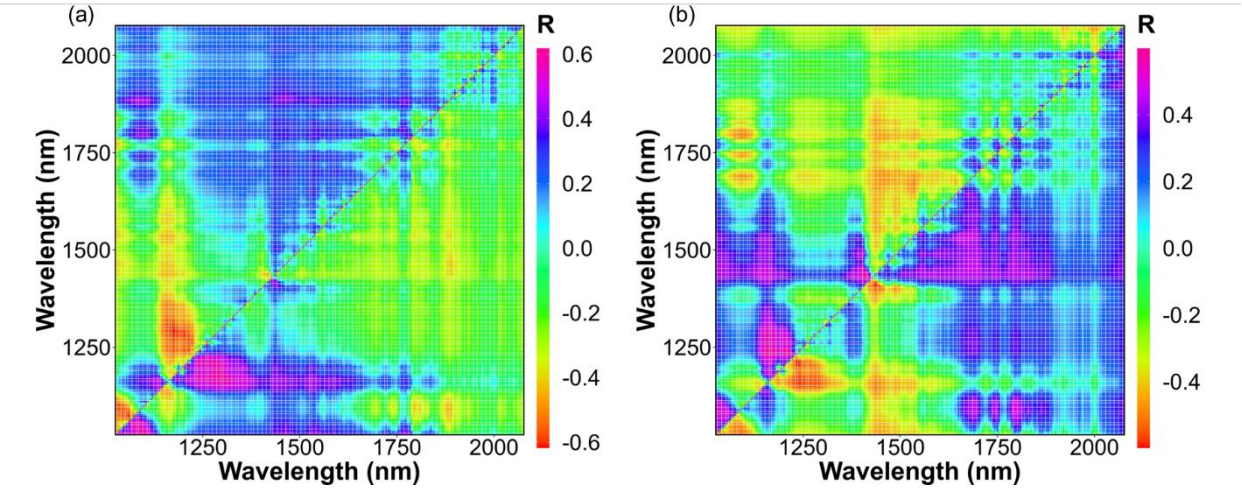

Figure 5.5: Correlation coefficient between NDSI and TSS (a) and TA (b)

5.3.3 Predictive Model Performance Based on Regression Models

The results obtained for testing the dataset based on different predictive models are shown in Table 5.3. Previous studies commonly used RMSE and R^2 to evaluate model performance, indicating that the best performance model show low RMSE and high R^2 values. In addition to RMSE and R^2, we used RPD and CCC to evaluate the model performance. RPD represents the ratio of a variable's standard deviation to its standard error of prediction using a specific model. For the RPD, it is considered as an almost perfect model when RPD > 2, moderate model when 2 > RPD > 1.4, poor model when RPD < 1.4 (Chang et al., 2001). CCC represents

how close a model prediction confirms to the ground truth data along a 45-degree line from the origin. For the CCC, the excellent mode with CCC > 0.8, moderate model with 0.8 > CCC > 0.65, poor model with CCC < 0.65 (Zhao et al., 2022).

For predicting TSS based on MSC+SG spectral pre-processing, the PLSR presented a RMSE, R^2, RPD and CCC of 0.554 °Brix, 0.583, 1.54 and 0.693 respectively. The result of predicting TSS based on the NDSI spectral pre-processing were a RMSE of 0.726 °Brix, and a R^2 of 0.305, a RPD of 1.114, and a CCC of 0.548 in the PLSR model. When it turns to TA, the PLSR showed a poor performance on both MSC+SG and NDSI pre-processing method (Table 5.3). The PLSR has been established as an effective predictive model to process hyperspectral data (Chen et al., 2015; Gomes et al., 2017a; Grafton et al., 2019). In this study, the result showed the PLSR had a moderate performance when predicting grape TSS through using MSC + SG pre-processing, while it had a poor performance when predicting grape TA. In addition to traditional regression models, we also explore the predictive power of two machine learning models (SVR and RFR). For SVR, the result of predicting TSS based on MSC+SG pre-processing showed the best prediction performance among all the regression models, with a result of RMSE, R^2, RPD and CCC of 0.523 °Brix, 0.622, 1.631, and 0.763, respectively. When using the selected NDSI as the input variables, the prediction performance of the SVR model decreased, but still showed moderate prediction performance (Table 5.3). For predicting TA based on the MSC+SG pre-processing features, the SVR presented a poor model performance with RMSE, R^2, RPD and CCC of 0.333 %, 0.03, 0.751, and 0.168, respectively. The combination of SVR and NDSI pre-processing showed a moderate prediction accuracy with RMSE, R^2, RPD and CCC of 0.19 %, 0.525, 1.463, and 0.675, respectively. For RFR, the result of predicting TSS based on MSC+SG pre-processing showed a poor model performance (RMSE = 0.642 °Brix, R^2 = 0.449, RPD = 1.368, CCC = 0.549). When the selected NDSI was used as input variables, the performance of RFR model was higher than that of using MSC+SG and showed a moderate performance (Table 5.3). When using MSC +SG as input variables, the RFR model showed the best model performance (RMSE = 0.217 %, R^2 = 0.43, RPD = 1.346, CCC = 0.595) to predict grape TA compared with other regression models (Table 5.3).

In addition, the combination of RFR and NDSI showed a moderate prediction performance to predict grape TA (Table 5.3). Figure 5.6 show the best model performance when predicting grape TSS and TA.

Table 5.3: Model performance on testing set based on regression models

Parameter	Feature	Model	RMSE	RPD	R^2	CCC
TSS (°Brix)	MSC+SG	PLSR	0.554	1.54	0.583	0.693
	MSC+SG	SVR	0.523	1.631	0.622	0.763
	MSC+SG	RFR	0.642	1.368	0.449	0.549
	NDSI	PLSR	0.726	1.114	0.305	0.548
	NDSI	SVR	0.543	1.49	0.589	0.655
	NDSI	RFR	0.592	1.429	0.531	0.687
TA (%)	MSC+SG	PLSR	0.305	0.82	0.025	0.15
	MSC+SG	SVR	0.333	0.751	0.03	0.168
	MSC+SG	RFR	0.217	1.346	0.43	0.595
	NDSI	PLSR	0.207	1.343	0.51	0.54
	NDSI	SVR	0.19	1.463	0.525	0.675
	NDSI	RFR	0.198	1.444	0.5	0.652

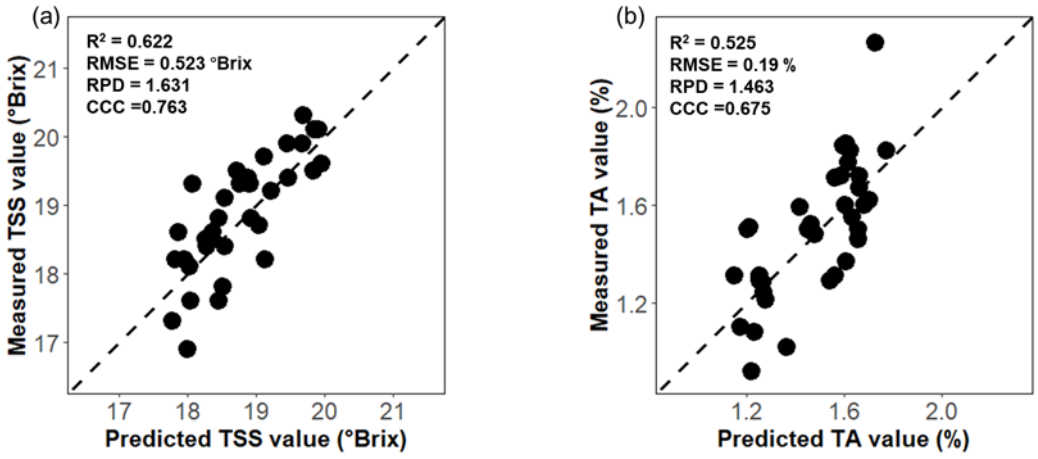

Figure 5.6: Predicted vs. measured values of TSS content of grapes using SVR and MSC + SG (a); Predicted vs. measured values of TA content of grapes using SVR and NDSI. The dot-ted line is the 1:1 line

5.3.4 Spectral Features Importance of TSS and TA

In order to identify the spectral signatures that contribute most to the model's performance, the feature importance of the best performance model based on the MSC+SG pre-processing was calculated (Figure 5.7). The PFI measures a feature's importance by calculating the increase of the model's prediction error after permuting the feature. For example, the reflectance value in band 1104 is most important for the SVR model predicting TSS, if permuting the wavelength 1104 nm in the data, the RMSE of the SVR model increase by a factor of around 1.04. In addition, PFI values showed that the electromagnetic spectrum contributed more at around 1088 – 1127, 1278 – 1322, 1810 -1838 nm (Figure 7a). The PFI values for the RFR model predicting TA fluctuated around at 1088 – 1104, 1138 – 1155, 1227, 1250, 1788 nm (Figure 5.7b). Figure 5.7 indicated that the PFI value are large at approximately 1100 nm in the TSS and TA prediction model. The peaks at 1100 nm were due to the vibrations related to the O-H and C-H bonds (Golic et al., 2003). It is worth noting that the NDSI values of the combination of these bands also have a high Pearson's correlation coefficient with TSS or TA.

Therefore, future studies should continue to explore the potential of using these relevant narrowband NDSI to classify or predict grape maturity.

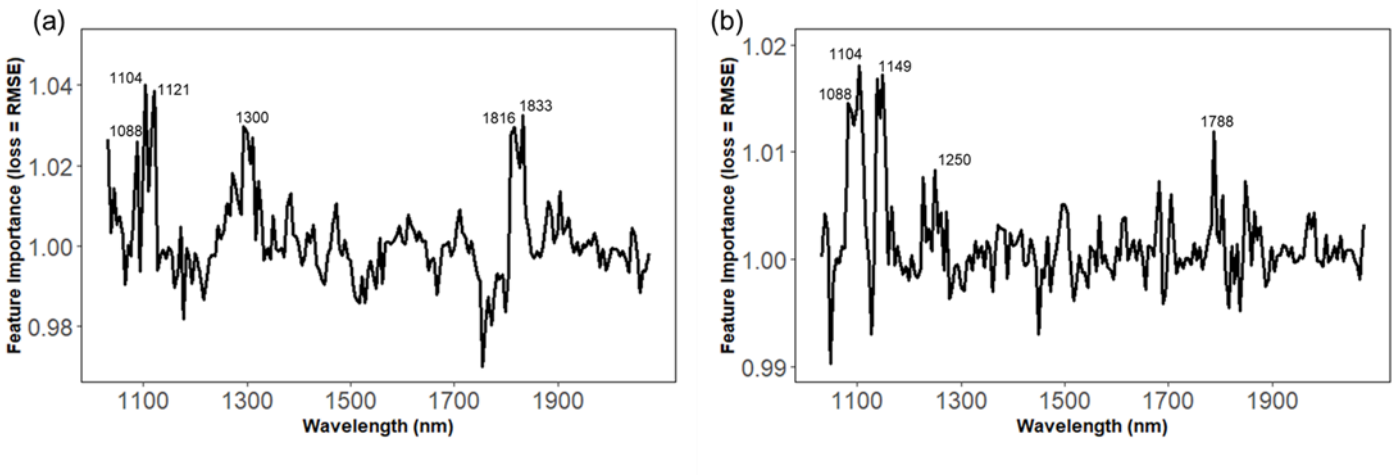

Figure 5.7: The permutation feature importance for TSS based on SVR model (a), TA based on the RFR model (b)

5.3.5 Discrimination Capacity Based on Classification Models

The classification analyses were performed for fresh berries using LDA, SVM and RF. The samples were divided into two groups (high and low quality) for each parameter and separated by threshold limiting values: 19 °Brix for TSS and 1.5% for TA. Table 5.4 showed the discrimination capacity of grape quality using different classification models and spectral pre-processing methods. The LDA model of TSS presented moderate classification capacity, with accuracy rate of 75% for MSC+SG dataset, and 68.8% for NDSI dataset. When using SVM discriminate TSS, 78.1% of samples were correctly classified with MSC+SG features, and 93.8% of samples were correctly classified with NDSI features. When using RF discriminate TSS, the percentage of total samples correctly classified was 84.4% with MSC+SG features, and 90.6% with NDSI features. In respect of the TA classification of fresh berries, the two quality groups were correctly classified with an accuracy rate from 75% to 84.4%. Compared with MSC+SG, the NDSI pre-processing has the ability to improve the classification performance when discriminating grape quality parameters.

Table 5.4: Discrimination performance on the testing set based on classification models

Parameter	Feature	Model	Acc	FP	Recall	Ce
TSS(°Brix)	MSC+SG	LDA	0.75	1	0.588	0.25
	MSC+SG	SVM	0.781	5	0.846	0.219
	MSC+SG	RF	0.844	2	0.8	0.156
	NDSI	LDA	0.688	7	0.769	0.313
	NDSI	SVM	0.938	1	0.933	0.0625
	NDSI	RF	0.906	2	0.933	0.094
TA (%)	MSC+SG	LDA	0.75	7	0.933	0.25
	MSC+SG	SVM	0.781	1	0.667	0.219
	MSC+SG	RF	0.844	2	0.8	0.156
	NDSI	LDA	0.75	7	0.941	0.25
	NDSI	SVM	0.844	3	0.857	0.156
	NDSI	RF	0.844	3	0.857	0.156

5.4 Discussion

The PLSR, SVR, and RFR models developed based on SWIR hyperspectral imaging system showed overall moderate performance for predicting TSS. When comparing the present work to previous work done in reflectance models, for TSS the present R^2 were slightly worse

than those of (Gomes et al., 2017a; Gutiérrez et al., 2019; Silva et al., 2018). One possible reason is that the range of spectrum used in previous study is between 400 – 1000 nm. However, when using RMSE as the model performance criteria, the present work showed higher performance ability than those of previous studies. We used RPD and CCC to further assess the regression model performance. The result showed that the combination of SVR and MSC + SG show a moderate prediction performance (RPD = 1.631, CCC = 0.763). The comparison done suggests that the spectrum ranged between 400 - 1000 nm may be more informative for quantifying the TSS value of grape berries. This is because a VIS - NIR HSI system to predict grape TSS value, maybe more practical, given its usually cheaper and has higher predictive performance than a SWIR HSI system for industry use. In addition, using a different variety of grape than previous work may have contributed to the poor performance of our model. Compared to TSS, the prediction for TA was lower in the regression models. The model accuracy was also worse than previous findings (González-Caballero et al., 2011; Guidetti et al., 2010). Previous studies have mainly used HSI between 400 -1000 nm to predict TA. (Guidetti et al., 2010) used a portable spectrophotometer to measure the oenological parameters in the spectral range 400 – 1000 nm. R^2 of 0.81 were obtained for estimating grape TA value in their study (Guidetti et al., 2010). The reduced predictive accuracy between the present work may be due to the different regression models, spectroscopic modes, and grape varieties. Future studies should continue to explore using current HSI systems to predict oenological parameters based on different grape varieties. It is noticeable that this study used two spectra pre-processing method: MSC+SG and NDSI The NDSI pre-processing method significantly improves the model prediction performance when predicting the TA value. Additionally, the smoothing techniques used in this study filter the noise of raw reflectance data possibly through the rapid exposure time. However, the smoothing technique will lose some characteristic values when processing spectral data, so further study should try to adjust camera parameters to avoid using smoothing techniques.

The study also examined the possibility of using HSI system to classify grape maturity. The SVM and RF models showed good classification performance for TSS. Compared with TSS,

the LDA, SVM and RF models only showed moderate classification performance for TA. The classification accuracy were higher than those of (Guidetti et al., 2010), who used the combination of spectrometer (1600 - 2400 nm) and PLS-DA to classify the grape maturity, and obtained 77.1% for TSS and 68.6% for TA. However, the 84.4% accuracy rate when classifying TA cannot meet the requirements of industrial application. The low sample size used in this study may have limited the classification model performance when classifying TA. Thus, further studies can use a large sample size with a wide range of TA value to examine the classification model performance. When compared with a study using spectrum in 350-2500 nm, our classification accuracy is similar to (dos Santos Costa et al., 2019) who used different supervised classification methods to determine the maturation stage of wine grapes, obtaining the accuracy rate from 92.26 to 93.15%. However, it is worth noting that running HSI or spectrometers in 400-2500 nm is expensive for industrial applications. Further research can continue to explore the possibility of selecting a specific band range for grading grape quality. In conclusion, classification models in this study would combine with current optical berry sorting systems to give winemakers a non-destructive tool for selecting the preferred berries.

5.5 Conclusion

The presented study explored the feasibility of applications of non-destructive means of estimating TSS and TA values of wine grapes by using SWIR, hyperspectral imaging with a spectral range of 1000 – 2500 nm. The Pearson's correlation coefficient between NDSI, TSS and TA showed that NDSI $(R_{1289}-R_{1177})/(R_{1289}+R_{1177})$ and $(R_{1411}-R_{1434})/(R_{1411}+R_{1434})$ which had a high correlation with TSS and TA, respectively. Furthermore, the PLSR, SVR, and RFR models were constructed to predict TSS and TA based on MSC+SG features or NDSI features. The prediction results demonstrated that the TSS of wine grapes can be predicted with moderate performance using SVR combined with an MSC+SG spectral pre-treatment process (RMSE = 0.523 °Brix, R^2 = 0.622, RPD = 1.631, and CCC = 0.763). Compared to TSS, the regression model performance for TA were relatively low, with the best model performance obtained from SVR combined with NDSI pre-processing (RMSE = 0.19%, R^2 = 0.525, RPD = 1.463, and CCC

= 0.675). We explored spectral feature importance of TSS and TA, and the result indicated that the significant regions and wavelengths relevant to TSS and TA estimation were at around 1100 nm. Finally, the LDA, SVM and RF models were performed to discriminate grape TSS and TA through different spectral pre-processing methods. Samples were divided into high or low quality for each parameter based on a threshold limiting value. The results showed that the best model correctly classify 93.8% of the samples for distinguishing TSS, and 84.4% of the samples for distinguishing TA. To enhance the industrial applicability of this technique, it is crucial to validate the developed regression or classification models on a broader range of wine grape varieties. This study has demonstrated the potential of using NIR-SWIR hyperspectral imaging for the rapid and non-destructive determination of TSS and TA value of wine grapes.

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Hongyi Lyu
Name and title of main supervisor:	Dr Miles Grafton, Senior Lecture
In which chapter is the manuscript/published work?	Chapter 6

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]

The candidate used hyperspectral imaging system to classify grape maturity in lab conditions based on deep learning models. Additionally deep learning model used to generate synthetic hyperspectral reflectance data of different grape maturity class.
The supervisory panel assisted with a hyperspectral spectrometer and other technical support, editing and writing support.

Please select one of the following three options:

○ The manuscript/published work is published or in press

Please provide the full reference of the research output:

◉ The manuscript is currently under review for publication

Please provide the name of the journal:
Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2024). Synthetic Hyperspectral Reflectance Data Augmentation by Generative Adversarial Network to Enhance Grape Maturity Classification.Postharvest Biology and Technology. (with editor)

○ It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	Hongyi Lyu 2024. 06. 29	Main supervisor's signature:	Miles Grafton Digitally signed by: Miles Grafton DN: CN = Miles Grafton email = m.grafton@massey.ac.bz C = NZ O = Massey University OU = SAE Date: 2024.06.29 14:02:03 +12'00'

This form should be placed at the beginning of each relevant thesis chapter.

[1] Refer to the Massey University Publishing and Authorship guidelines (OneMassey for staff, Stream for students) and/ or Contributor Roles Taxonomy (CRediT) guidelines for guidance.

6 Synthetic Hyperspectral Reflectance Data Augmentation by Generative Adversarial Network to Enhance Grape Maturity Classification

The previous chapter show moderate classification performance when using HSI under SWIR spectrum. However, the cost of SWIR hyperspectral cameras is higher than that of VNIR hyperspectral cameras. In this chapter, we use VNIR HSI combined with deep learning networks to determine the grape maturity. Considering the limited HSI data on grape berries, we propose a generative adversarial network to generate synthetic hyperspectral reflectance data to improve deep learning model performance.

This chapter is based on

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., & Sandoval, E. (2024). Synthetic Hyperspectral Reflectance Data Augmentation by Generative Adversarial Network to Enhance Grape Maturity Classification. Postharvest Biology and Technology (under review).

Abstract

Non-destructive and rapid grape maturity detection is important for the wine industry. The ongoing development of hyperspectral imaging techniques and deep learning methods has greatly helped in non-destructive assessing of grape quality and maturity, but the performance of deep learning methods depends on the volume and the quality of labeled data for training. Building non-destructive grape quality or maturity testing datasets requires damaging grapes for chemical analysis to produce labels which is time consuming and resource intensive. To solve this problem, this study proposed a conditional Wasserstain Generative Adversarial Network (WGAN) with the gradient penalty data augmentation technique to generate synthetic hyperspectral reflectance data of each grape maturity category. Thereafter, contextual deep three-dimensional CNN (3D-CNN) and spatial residual network (SSRN) are trained on synthetic and original datasets to classify grape maturity. After training of 10,000 epochs, synthetic hyperspectral reflectance data were very similar to real spectra for each maturity category. A SSRN model with synthetic hyperspectral reflectance data have shown the best classification performance with an average accuracy of 92%. The results show that both SSRN and 3D-CNN show better classification performance after adding synthetic hyperspectral reflectance data than using only an original dataset. This study indicated that deep learning models combined with conditional WGAN with the gradient penalty data augmentation technique had a good application prospect in the grape maturity assessment.

Keywords: hyperspectral imaging system; grape maturity; generative adversarial network; deep learning

6.1 Introduction

New Zealand wineries producing high-quality wines that are exported globally. One critical factor in the production of high-quality wines is the maturity of the grapes used to make the wine. Bramley et al., (2003) showed that it's more profitable for wineries to classify the maturity of grapes to produce wines of different qualities than to mix grapes of different maturity to produce a single quality wine. Traditional methods for measuring grape maturity involves measuring grape total soluble solids (TSS) based on refractometry or enzymatic tests (Rolle et al., 2022). These methods can provide accurate measurements, but are often labor-intensive, time-consuming, and require destructive sampling. However, the grape TSS value showed marked spatial variability within a block or in the same vine (Baluja et al., 2013). Using traditional destructive techniques only analyzed a small subset of grapes TSS values and use this value to represent the grape maturity of an entire region. This results in uneven maturity of the grapes used for winemaking during the post-harvest stage. Therefore, it is desirable to develop a non-destructive method to effectively sort grape maturity during the post-harvest stage.

The development of image technologies, e.g., monochromatic, red-green-blue (RGB), multispectral and hyperspectral imaging for non-destructive assessing of fruit quality or maturity has advanced rapidly over the last decades (Li et al., 2019; Lu et al., 2020, 2022). Hyperspectral imaging captures and processes information across a wide spectrum of wavelengths, allowing detailed analysis of various chemical compositions in grape berries (Lyu et al., 2024). Compared with other image technologies, hyperspectral imaging can detect subtle changes in the spectral reflectance of grape berries, which are related to their TSS value. Integrated with image technologies, deep learning improved efficacy and efficiency in grading and sorting grape berries. Deep learning, especially convolutional neural networks (CNNs) are widely being used in data-driven visual recognition (e.g., image classification and semantic segmentation) tasks in agricultural production (Wang et al., 2021). Various architectures of CNNs such as AlexNet, VGG, GoogLeNet, DenseNet, and Inceptionv3, are being used to classify fruit maturity status (Z. Zhao et al., 2023). Das and Yadav, (2020) proposed a revised version of AlexNet to

classify tomato maturity. Ashtiani et al., (2021) compared the accuracy of five different CNNs models to classify the maturity of white and black mulberries. The result showed that AlexNet and ResNet-18 models exhibited the best accuracy score for white and black mulberries, respectively. Additionally, Ramos et al., (2021) used VGG-19 to classify the maturity status of Syrah and Cabernet Sauvignon grapes. The result showed that VGG-19 achieved maturity status classification accuracy of 91.30% and 80.97% for Syrah and Cabernet Sauvignon, respectively.

One of the challenges in using deep learning for visual recognition tasks is the need for extensive and diverse datasets to train CNNs models effectively (Abbas et al., 2021). Deep learning models usually need train several million parameters to capture sufficient variations needed for vision recognition tasks (Miranda et al., 2023). In this case, deep learning models only have high performance and generalizability on sufficiently large and diverse datasets. For example, Zhao et al., (2023) used 3849 images from a “Nine Peach” dataset to train different segmentation models to determine peach maturity. They proposed a new one-stage instance segmentation model achieved an average precision of 72.12% on the large and highly diverse dataset. However, creating extensive and diverse datasets needs tremendous effort to ground truth data collection and labelling which is time consuming and resource intensive (Lu and Young, 2020). This is especially relevant for non-destructive fruit quality or maturity testing datasets, Labeling often necessitates damaging the fruit to measure internal qualities. To mitigate the scarcity of datasets, one common approach is data augmentation (Gour et al., 2020; Khalifa et al., 2022).

There are two categories of data augmentation approaches, including traditional image augmentation and augmentation based on deep learning methods. The traditional image augmentation include geometric (e.g., flipping and cropping) and color processing approaches (Shorten and Khoshgoftaar, 2019). However, methods such as cropping remove critical context from an image, making it harder for the model to learn the relationships between different parts of the image. Basic augmentation techniques often follow predictable patterns (e.g., fixed rotation angles, specific crops). This can lead to a lack of diversity in the training data, limiting

the model's exposure to truly novel examples. Additionally, traditional image augmentation can only create limited variations of existing data and is unable to learn the variations or invariant features across the samples (Chen et al., 2024). Recently, Generative Adversarial Networks (GANs) have been used for data augmentation to solve the drawback of traditional approaches in the agricultural community. The Generative Adversarial Network (GAN), originally proposed by Goodfellow et al., (2014) was used to synthesize new data with the same characteristics of training instances. GANs generate high-quality, realistic synthetic data that retain the statistical properties of the original dataset, enhancing the diversity without compromising the data quality (Shorten and Khoshgoftaar, 2019). Additionally, GANs introduce more diversity in the augmented data, capturing variations that basic image augmentation might miss.

Several studies have been used GANs as a data augmentation technique to synthesize fruit images in post-harvest quality assessment tasks (Guo et al., 2021; Bird et al., 2022; Miranda et al., 2023; Tan et al., 2024). For example, Bird et al., (2022) used Conditional GAN to produce synthetic lemon images for fruit quality image classification. When using synthetic fruit images in VGG-16 models, the classification performance is higher than that without data augmentation. However, most of the previous studies applied GANs to generate synthetic RGB or monochromatic fruit images. Currently, the potential of GANs in fruit hyperspectral image generation remains largely unexplored. Compared with RGB image, hyperspectral imaging can provide detailed spectral information for accurate substance identification, classification, and non-destructive internal quality assessment (Sun, 2010).

This study aims to develop a deep learning pipeline using hyperspectral imaging system to classify grape maturity in a non-destructive way. The specific objectives of this research are:

- To explore different CNNs models for grape maturity classification.
- To implement of a GAN for the generation of synthetic ripe and unripe grape hyperspectral reflectance data.
- To augment the original dataset with synthetic datasets to improve CNNs performance.

6.2 Methodology

6.2.1 Data Collection

The samples used in this study were collected from Palliser Estate Winery located in Martinborough. The winegrape variety used in this study was ‘Pinot Noir’, and compared to other varieties, the maturity status during the harvest stage greatly affects the value of subsequent wines. The ‘Pinot Noir’ were planted in an organic vineyard named Hua Nui. The vines were planted in 1998, with vine spacing of 1.7m and row spacing of 2.2m. A total of 231 samples were collected from a 2.6 ha block in Hua Nui vineyard (Figure 6.1). A total of 77 samples were collected, weekly from March 01 to March 14, 2024, during the harvest stage. To acquire representative samples, the sampling locations were collected from a predetermined nested grid. Each grid had locations taken at intervals of 4m, 8m, 16m and 32m. During each measurement, a healthy vine closest to the sampling location was selected for sampling. In each sampling vine, three bunches were selected for grape berry collection. The selected bunches were divided into three positions: top, middle, and bottom. From each position, one berry was handpicked resulting in a total of nine berries per sampling vine. It is worth noting that nine grape berries per vine met the minimum criteria for measuring TSS. The nine grape berries were bundled in plastic bags with labels and immediate processing spectrum measurements were taken.

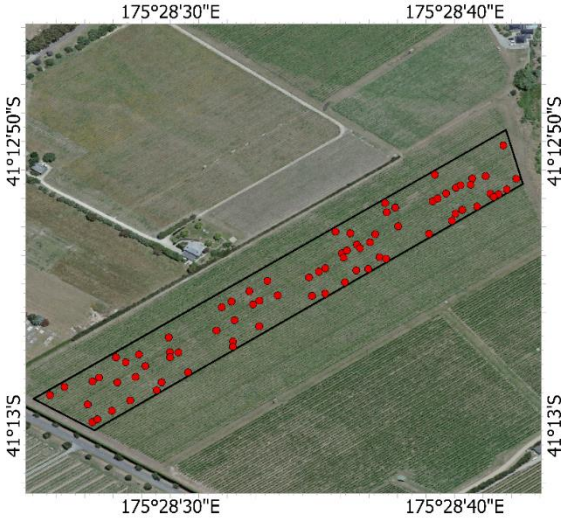

Figure 6.1: Sampling location in Hua Nui vineyard (red points represent each sampling point)

6.2.2 Hyperspectral Image Acquisition

The reflectance data of nine grapes were acquired from a push-broom VNIR hyperspectral imaging system with a wavelength from 406.8 to 995.8 nm. The VNIR hyperspectral imaging system is composed of the following constituents: two 150 W illumination bulbs, a spectra camera (HySpex VNIR-1800, Norsk Elektro Optikk, Norway), laboratory rack, white reference tile (SphereOptics GmbH, Germany), and a computer. The hyperspectral imaging system was turned on for 30 min before the image acquisition to increase the spectral data quality. The height between the camera and samples was 20.4 cm. HySpex Ground v4 12.7 software was used to process image acquisition within a dark room at room temperature. The frame period was 11200 us, integration time was 11085 us, and the moving speed of the conveyor belt is 5 cm/s. After acquiring hyperspectral images, the software HySpex Rad v3.4 was used to obtain the reflectance data of the image. Subsequently, the region of interest (ROI) tool was used in the Environment for Visualizing Images (ENVI) V 5.6 software (Research Systems Inc., Boulder, Co, USA) to manually separate the grape berries from the background.

6.2.3 Determination of Total Soluble Solids

After hyperspectral image acquisition, the samples were stored in the fridge at 4°C until their analysis for grape total soluble solids. Before each measurement, the samples were stabilized to the room temperature at around 25°C. During each measurement, nine grape berries were squeezed together into juice Then, 0.2 ml of the juice was applied onto a portable digital refractometer (PAL-BX/ACID F5 Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan). The TSS value (°Brix) was a direct reading from the refractometer with a measuring range between 0 and 60% (0.1% resolution). For each sample, the TSS value was measured twice. The average of these two measurements was used to represent the TSS value of the nine grapes.

6.2.4 Synthetic Image Generation via GAN Model

In order to improve the classification model performance, conditional Wasserstein GAN (WGAN) with the gradient penalty was used as the hyperspectral imaging data augmentation

technique to increase the size of the dataset. Traditional GAN consists of a generator and discriminator model which are trained in opposition to each other (Goodfellow et al., 2014). The work of the generator is to take random noise data as inputs and produce fake data. The work of discriminator is to receive the fake image from the generator and real data and distinguish whether the data are synthetic or real. The generator and discriminator train simultaneously and try to outdo each other. The discriminator makes sure that the data generated by the generator are as close to the real images as possible. The generator aims to produce data that the discriminator cannot distinguish from real data, while the discriminator aims to improve its accuracy in distinguishing real from synthetic data. This adversarial process ensures that the data generated by the generator become increasingly similar to the real images.

Compared with traditional GAN, WGAN use Wasserstein distance metric as its loss function. In WGAN, the discriminator is replaced by 'critic' science, it evaluates the quality of generated data by assigning them a score rather than classifying them as real or fake (Gulrajani et al., 2017). Conditional WGAN extends WGAN by incorporating additional information (e.g., class labels) into the network. Both the generator and critic are conditioned on this additional information, allowing the generator to produce class-specific data. This study introduces an additional classifier model in the network (Figure 6.2). The classifier ensures that the generated data not only appears realistic but also adheres to the given conditional labels (Audebert et al., 2018). This setup helps in aligning the generated data distribution more closely with the conditional labels provided. The objective function for the generator and critic, including the classifier, is:

$$\min_{G} \max_{D} (E_{x\sim P_{data}}[D(x,c)] - E_{z\sim P_z, c\sim P_c}[D(G(z,c),c)] + \lambda E_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x},c)\|_2 - 1)^2])$$

where x is real data, P_{data} is the data distribution, $D(x,c)$ is the critic's score for the real data x conditioned on c. z is noise sample, P_z is the noise distribution, c is conditional label, P_c is label distribution, $G(z,c)$ is fake data from the generator using noise z and condition c. $D(G(z,c),c)$ is critic's score for the fake data $G(z,c)$ conditioned on c. λ is coefficient for the gradient penalty term, $\hat{x}$ are samples interpolated between real data x and fake data $G(z,c)$, P_x

is the distribution of these interpolated samples, $D(\hat{x}, c)$ is critic's score for interpolated samples $\hat{x}$ conditioned on c.

The conditional WGAN with the gradient penalty generator model consists of several dense layers followed by Leaky rectified linear unit (ReLU) activation functions and a final Sigmoid activation function. The model begins with an input layer that takes in the noise sample and condition vectors. This is followed by a sequence of three dense layers, each followed by a Leaky ReLU activation layer. The final dense layer is followed by a Sigmoid activation layer to produce smooth outputs in the range [0,1]. The critic and classifier model in the conditional WGAN with the gradient penalty have similar architecture, including an input layer, three dense layers followed by Leaky ReLU activation functions. The final dense layer in critic model outputs a single value, which indicates the authenticity of the input data. The final dense layer in the classifier model outputs a vector with a size equal to the number of classes, which represents the classification scores for each class. The model applies a weight initialization function to all layers, following the He et al., (2016) policy, which initializes weights using the Kaiming normal distribution.

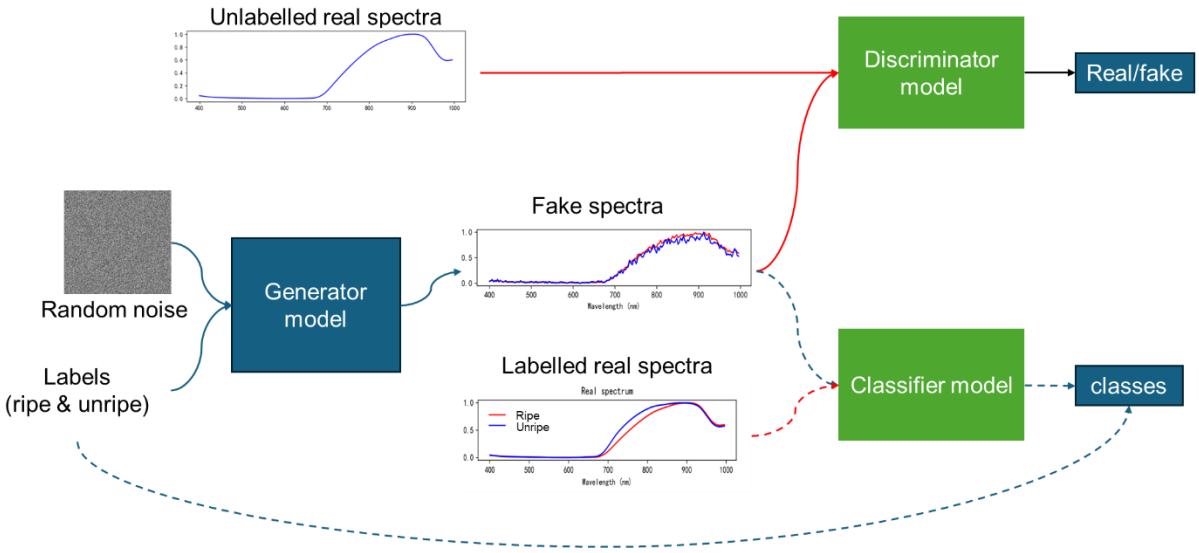

Figure 6.2: The proposed conditional WGAN with the gradient penalty. Red lines represent the training of discriminator; blue lines represent the training of the generator; red and blue dashed lines represent the training of classifier

6.2.5 Grape Maturity Classification with Different CNNs Models

As this study aimed to assess the effectiveness of GAN-based data augmentation methods in enhancing grape maturity classification performance, two widely used hyperspectral imaging classification CNN models: contextual deep three-dimensional CNN (3D-CNN) and spatial residual network (SSRN) were used. 3D-CNN is an end-to-end fully convolutional network proposed by Lee and Kwon, (2017). The 3D-CNN uses a multi-scale convolutional filter bank. This multi-scale approach captures features at various spatial resolutions to analysis both spectral and spatial information in the hyperspectral data. Additionally, 3D-CNN incorporates residual learning modules to help the network learn deeper representations without encountering the vanishing gradient problem. The SSRN uses 3-D convolutional layers to process the hyperspectral data cubes, incorporating both spectral and spatial information effectively (Zhong et al., 2017). SSRN incorporates spectral and spatial residual blocks to improve training efficiency and performance. Spectral residual blocks use 1 × 1 × m convolutions, focusing on spectral features, while spatial residual blocks use convolutions that maintain spatial dimensions to extract spatial features. To regularize the training process and reduce overfitting, batch normalization and dropout techniques are applied within the network (Zhong et al., 2017). Appendix 3 shows hyperparameter settings for different CNN models.

6.2.6 Experiment Setup and Evaluation Metrics

The samples were divided into two groups (unripe and ripe) separated by threshold limiting values: 21 °Brix for TSS. In this study, two sets of experiments have been performed. First of all, the conditional WGAN with the gradient penalty has been trained for 10,000 epochs to generate synthetic hyperspectral grape reflectance data of two categories (unripe and ripe). After successful completion of the training of conditional WGAN with the gradient penalty, 3500 synthetic hyperspectral data of grape berries were generated for each category.

In the second step, the classification model (3D-CNN and SSRN) has been trained on the training set and synthetic dataset. In order to compare the performance of each classifier before and after data augmentation, the original training set was defined as a reference baseline.

Subsequently synthetic data is added to the original set. The samples were randomly split into training and testing sets using a 30/70 ratio. The split step was repeated three times with different random seeds in order to improve the stability of the model. The testing set was used to evaluate the performance based on the accuracy (Acc), recall, classification error Rate (Ce) and Precision which were calculated from the confusion matrices of each model. The calculation of Acc, recall, and Ce was defined bellow:

$$Acc = \frac{TP+TN}{TP+FN+FP+TN} \quad \text{Equation 6-1}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{Equation 6-2}$$

$$Ce = \frac{FP+FN}{TP+FN+FP+TN} \quad \text{Equation 6-3}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{Equation 6-4}$$

where TP, FN, FP, and TN are the true positive, false negative, false positive, true negative.

The model in this work were implemented with PyTorch (version 1.8) and CUDA 11.1 and trained on an RTX 2080Ti GPU (4352 CUDA cores).

6.3 Result

6.3.1 Basic Statistical Analysis of Measured TSS

The basic statistical description (i.e., sampling date, mean, maximum, minimum, stand deviation, coefficient of variation and variance) of the TSS value is shown in Table 6.1. The TSS value of grapes increases closer to the harvest date.

Table 6.1: Descriptive statistics for TSS value of Pinot Nor during different sampling date (where SD = stand deviation, CV = coefficient of variation)

Date	Mean	Maximum	Minimum	SD	CV	Variance
01 March	19.48	21.3	18	0.71	4%	0.5

06 March	20.67	23.1	19	0.83	4%	0.67
14 March	22.06	23.9	20	0.76	3%	0.57

6.3.2 Performance of Synthetic Hyperspectral Reflectance Data

After training from the conditional WGAN with the gradient penalty, this study generates 3500 hyperspectral reflectance data for each category (unripe and ripe). To evaluate the quality of synthetic data, Figure 6.3 shows comparison results of real spectrum and synthetic data spectrum for different maturity classes. Figure 6.3 shows that when the epoch = 500, the synthetic data spectrum was random noise. When epoch = 1,000, the synthetic data spectrum produces a spectral curve with large noise values. As the number of epochs grows to 5,000 epoch, the synthetic data spectrum was gradually smoothed, while the synthetic data spectrum cannot show the characteristics of the different maturity classes. Finally, it could be seen that after epoch = 8,000 the synthetic data spectrum of different maturity classes show difference in the Red Edge region, consistent with the real spectrum.

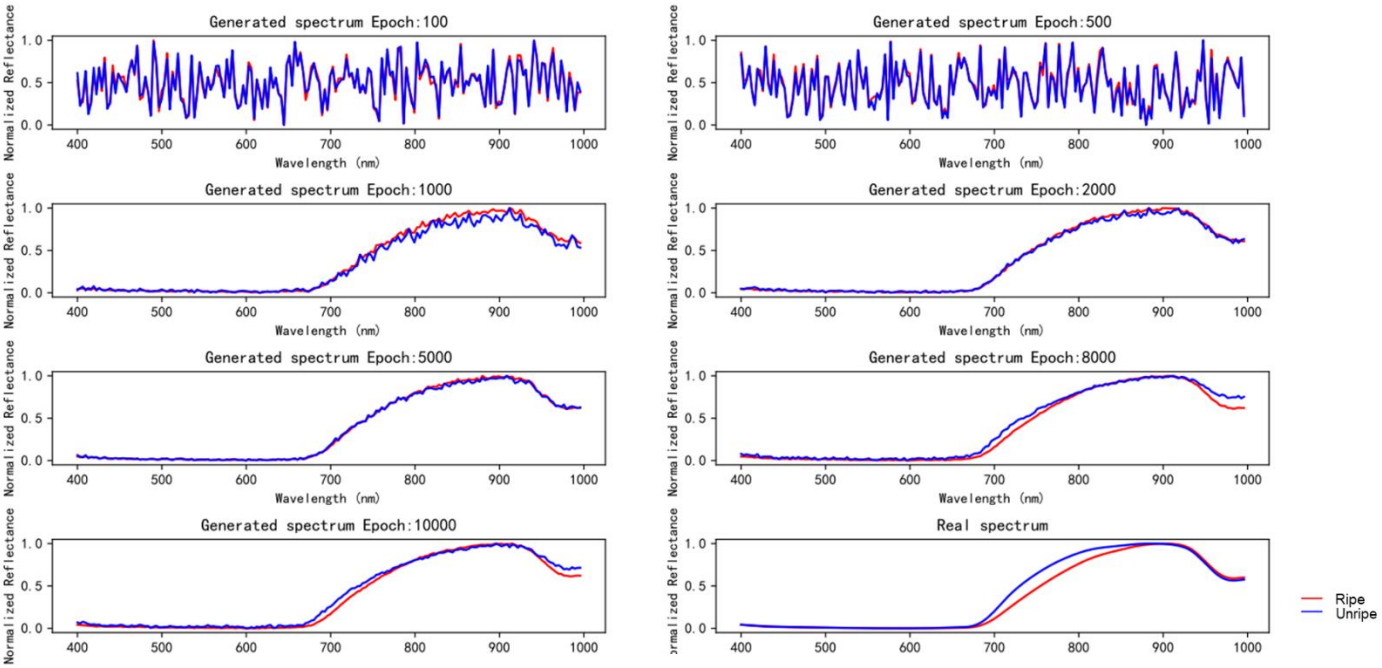

Figure 6.3: Spectrum generated by the conditional WGAN with the gradient penalty

6.3.3 Performance of Classification Model

This study evaluated the performance of proposed model for binary classification (unripe and ripe) task. Table 6.2 - 6.5 represent the classification performance of the proposed method on the original dataset and augmented dataset (original + synthetic data datasets) with different random seeds. The classification performance on the original dataset was set as the reference in order to compare the impact of GAN data augmentation techniques on model performance. In this study, two common hyperspectral imaging classification deep learning models (SSRN and 3D-CNN) were used to assess the GAN data augmentation techniques. Table 6.2 shows the SSRN model with synthetic hyperspectral reflectance data has the best classification performance with average accuracy of 92% on the test set. When using the SSRN model with the original dataset, the SSRN model achieved a classification accuracy of 87%, 89% and 90% for three random seed settings (Table 6.2). Additionally, the performance of using 3D-CNN with synthetic hyperspectral reflectance data outperformed predictions using the original dataset on the test set, with classification accuracy of 91%, 92% and 92% (Table 6.4). The performance of SSRN and 3D-CNN also was evaluated on the total dataset (Table 6.3, 6.5). The result show that both SSRN and 3D-CNN have better classification performance after adding synthetic hyperspectral reflectance data than using the original dataset only.

Table 6.2: Classification performance of SSRN on test set

	Original				Original + Synthetic			
	Acc	Recall	Ce	Precision	Acc	Recall	Ce	Precision
Seed 1	0.87	0.81	0.13	0.96	0.92	0.99	0.08	0.88
Seed 2	0.89	0.85	0.11	0.94	0.92	0.98	0.08	0.88
Seed 3	0.9	0.93	0.1	0.89	0.92	0.99	0.08	0.87

Table 6.3: Classification performance of SSRN on total set

	Original				Original + Synthetic			
	Acc	Recall	Ce	Precision	Acc	Recall	Ce	Precision
Seed 1	0.88	0.83	0.12	0.96	0.93	0.99	0.07	0.88
Seed 2	0.89	0.87	0.11	0.94	0.93	0.99	0.07	0.89
Seed 3	0.91	0.94	0.09	0.9	0.92	0.99	0.07	0.87

Table 6.4: Classification performance of 3D-CNN on test set

	Original				Original + Synthetic			
	Acc	Recall	Ce	Precision	Acc	Recall	Ce	Precision
Seed 1	0.84	0.79	0.16	0.92	0.91	0.93	0.09	0.9
Seed 2	0.81	0.75	0.19	0.9	0.92	0.99	0.08	0.88
Seed 3	0.84	0.82	0.16	0.9	0.92	0.99	0.08	0.88

Table 6.5: Classification performance of 3D-CNN on total set

	Original				Original + Synthetic			
	Acc	Recall	Ce	Precision	Acc	Recall	Ce	Precision
Seed 1	0.86	0.82	0.14	0.93	0.91	0.92	0.09	0.9
Seed 2	0.8	0.75	0.2	0.89	0.92	0.99	0.08	0.87
Seed 3	0.85	0.82	0.15	0.9	0.92	0.99	0.08	0.87

6.4 Discussion

Non-destructive assessing of grape maturity during the post-harvest stage is important to wineries to classify the grapes and produce wines of different quality. This study aims to use hyperspectral imaging system (400-1000 nm) and deep learning techniques to classify grape maturity in a non-destructive and rapid way. The TSS classification accuracy was moderate when SSRN was employed, where 3D-CNN had the best classification accuracy of 90% (Table 6.2 – 6.5). Most previous studies restricted their experiments to regression tasks to predict the value of grape TSS (Silva and Melo-Pinto, 2023; Tsakiridis et al., 2023). However, wineries are more concerned about the maturity grade of the grape than the specific TSS value. The classification task is more suitable for non-destructive assessing of grape maturity than the regression task. The classification result of this study support the finding of Ramos et al., (2021), who used VGG-19 To classify the maturity status of Syrah and Cabernet Sauvignon. The ability of using CNNs models to process big data is no longer the bottleneck, instead these methods require large datasets than are usually available or easily obtained.

In real practical conditions, building CNNs models need thousands of images or data to train the model to prevent the model from overfitting. However, collecting thousands of images and data for non-destructive grape maturity needs large resources to ground truth data collection and labelling which is time consuming (Lu and Young, 2020). To mitigate the scarcity of datasets, this study proposed a data augmentation method based on deep learning architecture. Conditional WGAN with the gradient penalty was used to produce synthetic hyperspectral reflectance data. Compared with traditional GANs architecture, this study introduces an additional classifier model in the network. The classifier improves the quality and diversity of the generated samples by conditioning them on specific class labels. To balance the trade-off between data quality and computational efficiency, the conditional WGAN with the gradient penalty has been trained for 10,000 epochs (Figure 6.3). The result show that the synthetic hyperspectral reflectance data generated by the conditional WGAN with the gradient

penalty are closely similar to real data spectrum for each maturity class. Most previous studies applied GANs to generate RGB images in the agricultural and food production field (Abbas et al., 2021; Bird et al., 2022; Lu et al., 2022). This study demonstrates the potential of using GANs to generate hyperspectral reflectance data. This has helped by generating a large and diverse hyperspectral reflectance dataset.

The results using SSRN and 3D-CNN with synthetic hyperspectral reflectance data can be observed from Table 6.2 to Table 6.5. Synthetic hyperspectral reflectance data plus the original dataset scored higher classification accuracy than training only on the original dataset, showing that augmentation has had a positive effect on the classification model performance. This is similar to the work of Abbas et al., (2021) who used conditional GAN to generate tomato plant disease images. They showed the deep learning models have higher classification accuracy after adding synthetic image data during the training process. Future work should continue to explore using GANs to generate different maturity classes of grapes under different spectral regions. The industrial use of cameras only focuses on a specific spectrum region, exploring the ability of using GANs in specific regions that can be better applied to industrial data generation. Additionally, the nature of datasets acquired from hyperspectral imaging systems include the background and grape berries. Hence, image segmentation is required to separate the berry from the background manually. If an end-to-end deep learning segmentation network is trained and used before our proposed approach, this will save labor costs and enable rapidly non-destructive assessing of grape maturity.

6.5 Conclusion

Deep learning models have shown promising results in fruit quality or maturity detection during the post-harvest stage. However, training a deep learning model needs extensive and diverse datasets to avoid overfitting. In addition, labelling the grape maturity status also needs destructive analysis, usually from a limited dataset. This study proposes conditional WGAN with the gradient penalty to generate synthetic hyperspectral reflectance data in order to provide an extensive and diverse grape maturity dataset. The performance of proposed data augmentation

techniques was evaluated on two CNNs (SSRN and 3D-CNN). Synthetic hyperspectral reflectance data plus original dataset scored higher classification accuracy than training only on the original dataset, showing that augmentation has had a positive effect on the classification model performance. The SSRN model with synthetic hyperspectral reflectance data has shown the best classification performance with average accuracy of 92%. This study demonstrates the potential of using GANs to generate hyperspectral reflectance data and improve the deep learning model classification accuracy. In the future, GANs are expected to be applied on more grape varieties and specific spectrum regions.

7 Overall Summary

In New Zealand, winegrape growers have traditionally used destructive techniques to assess grape quality parameters and determine the optimal harvest time. To establish the harvest date, grapes are sampled weekly from veraison onward to see how the levels of different grape quality parameters procede through the season. Typically, a small sample (around 200 berries per block and per variety) is used for this purpose. However, significant spatial variation in grape quality exists within a block and even within individual clusters. Traditional harvesting techniques can lead to uneven grape quality, adversely affecting the value of the resulting wines. The main aim of this thesis was to develop a suitable non-destructive technique to predict 'Pinot Noir' quality parameters at different stages of vine growth. So, as to assist winegrape growers in making more optimal harvesting decisions and achieving selective harvesting, thereby increasing profitability in the wine industry.

‘Pinot Noir’ winegrapes were sourced commercially from different vineyards in Martinborough, different harvesting seasons, and various berry growth stages. Three non-destructive techniques were evaluated: remote sensing imagery (Chapter 3), VNIR-SWIR spectroradiometer (Chapter 4), and hyperspectral imaging system (Chapter 5 and 6), in order to estimate grape quality parameters during pre - or post - harvest stages.

7.1 Integrating Data from Multiple Source and Machine Learning for Predicting Grape Quality in Vineyards

The goal of using remote sensing imagery and ancillary data to predict grape TSS is to provide accurate and large-scale estimates of grape maturity and quality. Predicting the quality of grapes during the pre-harvest stage can assist growers make harvesting schedule in advance and achieve a selective harvest. Chapter 3 evaluates the effectiveness of different vegetation indices (VIs) obtained from UAV multispectral imagery, alongside other ancillary variables, in predicting grape TSS using machine learning models. The Pearson's correlation coefficient between 23 VIs and grape TSS was calculated. The result showed a strong

correlation (r=0.64) between the OSAVI and grape TSS. Thus, the OSAVI was selected as the primary input variable with other ancillary data including: NDVI_proximal, soil ECa, elevation, slope from LiDAR data, trunk circumference and day of year to build regression models for estimating grape TSS. Additionally, NGBDI was selected as the primary input because it is available from cheaper RGB cameras and shows a strong correlation with grape TSS. Seven machine learning models were tested, ridge regression, lasso regression, k-Nearest Neighbor (KNN), support vector regression (SVR), random forest regression (RFR), extreme gradient boosting (XGBoost), and artificial neural networks (ANN). RFR was the best model for both indices, achieving slightly better performance with NGBDI (R^2=0.54, RMSE=1.16 °Brix) compared to OSAVI (R^2=0.51, RMSE=1.19 °Brix).

The findings of Chapter 3 highlight the potential of using remote sensing imagery combined with ancillary data for precision viticulture. The strong correlation between VIs and grape TSS suggests that UAV-based multispectral imaging can effectively monitor and predict grape quality across large vineyard areas. Additionally, machine learning models show the ability to incorporate data from multiple sources (e.g., UAV imagery, soil ECa, LiDAR data) enhances their predictive performance, offering a robust tool for vineyard management. Although the prediction performance in Chapter 3 is less than satisfactory, these can be linked to the limited sample berries that were used to represent the vine TSS status. The study's sampling method, which used only three berries per vine, may not accurately represent the TSS variability across the entire vine. Grape TSS is influenced by various environmental conditions during the growing season, including radiation and temperature. These factors can cause fluctuations in grape quality that are not fully captured by the VIs and ancillary variables used in the models. Further research should explore the inclusion of more dynamic environmental variables.

7.2 Integrating Spectral Data from VNIR-SWIR Spectroscopy and Machine Learning for Predicting Grape Quality in Vineyards

The research in relation to describing grape quality through VNIR-SWIR spectroscopy has been constrained to a specific location and under stable laboratory conditions. Thus, the

practical ability of VNIR-SWIR spectroscopy in field conditions was not known. In order to explore the potential of using VNIR-SWIR spectroscopy to predict grape quality under uncontrolled field conditions, an experiment was conducted (Chapter 4). Reflectance data of each grape berry in the spectral range from 350 to 2500 nm was acquired through a portable spectroradiometer (ASD FieldSpec 4 Hi-Res NG Spectroradiometer). The spectroradiometer provides a sampling interval of 1.4 nm from 350 to 1000 nm, and 1.1 nm from 1001 to 2500 nm. These intervals were then interpolated to 1 nm, resulting in 2151 values for every spectral measurement. The spectroradiometer was equipped with a contact probe for controlled illumination. After recording the berry reflectance data, the raw reflectance was processed by various transformations including spectral pre-processing and dimensionality reduction.

The relationship between spectral data and grape TSS were established through four machine learning models (KNN, RFR, SVR, and ES). ES model uses diverse learners to improve accuracy. By leveraging their differences, it provides complementary information and reduces the risk of incorrect results. The result on the independent test set showed that, the ES model based on first derivative spectra data, in the VNIR-SWIR region provided the highest prediction accuracy for grape TSS value, with a R^2 of 0.803, RMSE of 1.213 °Brix. This result has advanced the understanding, that the majority of grape TSS can be predicted by VNIR-SWIR spectroscopy within, uncontrolled field environmental conditions. Additionally, Chapter 4 examines the effect of different spectral regions (VNIR, SWIR, VNIR-SWIR) on predicting grape TSS. The result suggests that using VNIR equipment is sufficient for winegrape growers, as VNIR-SWIR did not significantly improve model performance but did require more expensive equipment costs. The results confirm the suitability of using VNIR-SWIR spectroscopy to predict ‘Pinot Noir’ TSS values under field conditions. This non-destructive and rapid method can help winegrape growers to better determine harvest strategies.

7.3 Integrating Spectral Data from Hyperspectral Imaging System and Machine Learning for Predicting Grape Quality

Although the results of Chapter 4 were satisfactory, VNIR-SWIR spectroscopy only can provide spectra information of grapes and cannot provide information on spatial resolution of traits in grape berries. Previous studies have been devoted to exploring using HSI in 400 – 1000 nm to predict grape quality parameters. However, studies related to using HSI in 1000 – 2500 nm to sort the grape quality have been rare. In order to fill this gap, a Specim SWIR HSI system was tested, by collecting reflectance data from grape berries (Chapter 5). A series of regression and classification models were applied to the reflectance data, in order to determine grape TSS and TA. Spectral pre-processing techniques were employed to enhance the quality of the spectral data including MSC+SG and NDSI.

The SVR model with MSC+SG preprocessing exhibited the best performance for predicting TSS, achieving an RMSE of 0.523 °Brix and an R^2 of 0.622. For TA prediction, the SVR model combined with NDSI preprocessing showed an RMSE of 0.19% and an R^2 of 0.525. The reason for the low accuracy of predicting grape TSS and TA appears to be due to the spectral range used in this study is in 1000 – 2500 nm. The wavelength which had the highest importance in the best performance model prediction were highlighted by using a permutation feature importance (PFI) method. PFI values were large at approximately 1100 nm in the TSS and TA prediction model. This wavelength could be used to build a specific multispectral sensor for predicting grape TSS and TA. The classification models demonstrated strong performance in distinguishing grape quality. The SVM model with MSC+SG preprocessing showed the best classification performance to classify grape TSS, with accuracy of 93.8%. The RF model with MSC+SG preprocessing showed the best classification performance to classify grape TA, with accuracy of 84.4%. The high classification accuracy of classifying grape TSS indicates the efficacy of using HSI and machine learning to build a gape TSS sorting system in order to help wine industry to achieve graded wine production.

7.4 Integrating Generative Adversarial Network for Generating Synthetic Hyperspectral Reflectance Data

In order to extend the potential of using HSI to determine the grape quality, it was further evaluated by another HSI within the VNIR spectrum region (Chapter 6). In this study deep learning networks was used to classify the grape TSS value. Compared with machine learning algorithms, deep learning networks provide a multi-layer approach, to capture complex patterns and hierarchical structures in high-dimensional data. However, the performance of deep learning methods depends on the volume and the quality of labeled data for training. Building non-destructive grape quality testing datasets requires damaging grapes for chemical analysis to produce labels, which is time consuming and resource intensive. To solve this problem, Chapter 6 proposed a conditional Wasserstain Generative Adversarial Network (WGAN) with the gradient penalty data augmentation technique to generate synthetic hyperspectral reflectance data of each category. The results show that both SSRN and 3D-CNN show better classification performance after adding synthetic hyperspectral reflectance data than using only an original dataset. Therefore, the methods used in Chapter 6 could be used to generate synthetic hyperspectral reflectance data of grapes and to extrapolate to more grape varieties and different spectrum regions. However, the labeled data features must be representative of the total data set features being generated.

7.5 Conclusion

The dissertation achieved several significant results, primarily focusing on the development and validation of non-destructive techniques to predict 'Pinot Noir' grape quality parameters. The main findings include:

- UAV based multispectral imagery combined with ancillary data showed a moderate prediction performance to predict grape TSS. A strong correlation between certain vegetation indices, particularly OSAVI and NGBDI, and grape TSS was found. Random

forest regression (RFR) emerged as the best predictive model when predicting TSS, achieving R^2=0.54 with NGBDI and R^2=0.51 with OSAVI.

- The use of VNIR-SWIR spectroscopy under field conditions was assessed. Enhanced prediction accuracy of grape TSS using stack ensemble learning model with first derivative spectra data, achieving R^2=0.803 and RMSE=1.213 °Brix. The spectroscopy under VNIR spectrum is sufficient for accurate prediction, suggesting cost-effectiveness over VNIR-SWIR.

- Investigated HSI in the 1000–2500 nm range to predict grape TSS and TA value. Achieved high classification accuracy for TSS (93.8%) and TA (84.4%) using support vector machine (SVM) and random forest (RF) models. Highlighted wavelengths (around 1100 nm) critical for grape quality prediction.

- Investigated HSI in the 400-1000 nm range to classify grape TSS value. Proposed a conditional Wasserstein GAN (WGAN) with gradient penalty for generating synthetic hyperspectral reflectance data in the VNIR spectrum region. Demonstrated improved classification performance of deep learning models (SSRN and 3D-CNN) with synthetic data, paving the way for broader application and dataset augmentation.

The findings of this dissertation provide vineyard managers with advanced tools and insights to optimize their decision-making processes:

- Non-destructive techniques allow for more accurate and real-time assessment of grape quality, aiding in the optimal timing of harvest to enhance wine quality.

- The ability to predict grape quality spatially across a vineyard enables selective harvesting, ensuring uniformity and high quality in the final wine product.

- Utilizing VNIR spectroscopy offers a cost-effective alternative to more expensive VNIR-SWIR systems without compromising prediction accuracy.

- Implementation of UAV-based remote sensing can cover large vineyard areas efficiently, reducing labor and operational costs.

- Integrating various data sources (e.g., multispectral imagery, soil ECa, LiDAR data) into machine learning models provides a robust decision-support tool.
- Enhanced prediction models and synthetic data generation techniques can lead to better resource allocation and strategic planning

7.6 Future Research

- Even though the prediction results of grape TSS using machine learning and deep learning models were satisfactory. The recently developed computational models (Transformer and Diffusion model) need to be explored as they may improve the prediction accuracy.
- UAV based multispectral imagery only provides moderate prediction results for estimating grape TSS, the study methods need to be improved by collecting more grape berries from the sample vine in order to better represent the quality status of the selected vine.
- Satellite based multispectral imagery need to be used in further study Satellite imagery is easier to get than UAV data. However, the spatial resolution of current open-source satellite data is mostly above 10m, which cannot meet the prediction of a single vine's yield and quality status. Future studies should improve satellite spatial resolution in conjunction with current super-resolution algorithms and explore its ability to predict vine quality and yield.
- Chapter 4 and Chapter 6 showed a satisfactory prediction and classification result when using VNIR-SWIR spectroscopy and HSI to determine grape TSS. However, this study only uses 'Pinot Noir' as the sample variety. Future studies should extend the application of developed techniques and algorithms to other grape varieties and regions to validate their generalizability and adaptability.

- Based on key wavelengths identified in Chapter 5, future studies should develop specific multispectral sensors tailored for grape quality prediction, enhancing practical deployment.

8 Reference

Abbas, A., Jain, S., Gour, M., Vankudothu, S., 2021. Tomato plant disease detection using transfer learning with C-GAN synthetic images. Computers and Electronics in Agriculture 187, 106279.

Agulheiro‐Santos, A.C., Ricardo‐Rodrigues, S., Laranjo, M., Melgão, C., Velázquez, R., 2022. Non‐destructive prediction of total soluble solids in strawberry using near infrared spectroscopy. Journal of the Science of Food and Agriculture 102, 4866–4872.

Albetis, J., Duthoit, S., Guttler, F., Jacquin, A., Goulard, M., Poilvé, H., Féret, J.-B., Dedieu, G., 2017. Detection of Flavescence dorée grapevine disease using unmanned aerial vehicle (UAV) multispectral imagery. Remote Sensing 9, 308.

Albetis, J., Jacquin, A., Goulard, M., Poilvé, H., Rousseau, J., Clenet, H., Dedieu, G., Duthoit, S., 2018. On the potentiality of UAV multispectral imagery to detect Flavescence dorée and Grapevine Trunk Diseases. Remote Sensing 11, 23.

Anastasiou, E., Balafoutis, A., Darra, N., Psiroukis, V., Biniari, A., Xanthopoulos, G., Fountas, S., 2018. Satellite and proximal sensing to estimate the yield and quality of table grapes. Agriculture 8, 94.

Arab, S.T., Ahamed, T., 2022. Land suitability analysis for potential vineyards extension in Afghanistan at regional scale using remote sensing datasets. Remote Sensing 14, 4450.

Arab, S.T., Noguchi, R., Matsushita, S., Ahamed, T., 2021. Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach. Remote Sensing Applications: Society and Environment 22, 100485.

Arnó, J., Rosell, J.R., Blanco, R., Ramos, M.C., Martínez-Casasnovas, J.A., 2012. Spatial variability in grape yield and quality influenced by soil and crop nutrition characteristics. Precision Agriculture 13, 393–410.

Arnó Satorra, J., Martínez Casasnovas, J.A., Ribes Dasi, M., Rosell Polo, J.R., 2009. Precision viticulture. Research topics, challenges and opportunities in site-specific vineyard management. Spanish Journal of Agricultural Research, 2009, vol. 7, núm. 4, p. 779-790.

Ashtiani, S.-H.M., Javanmardi, S., Jahanbanifard, M., Martynenko, A., Verbeek, F.J., 2021. Detection of mulberry ripeness stages using deep learning models. IEEE Access 9, 100380–100394.

Audebert, N., Le Saux, B., Lefèvre, S., 2018. Generative adversarial networks for realistic synthesis of hyperspectral samples, in: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 4359–4362.

Baggiolini, M., 1952. Les stades repères dans le développement de la vigne et leur utilisation pratique, Station Féd. Essais Agric., Lausanne.

Baluja, J., Tardaguila, J., Ayestaran, B., Diago, M.P., 2013. Spatial variability of grape composition in a Tempranillo (Vitis vinifera L.) vineyard over a 3-year survey. Precision Agriculture 14, 40–58.

Benelli, A., Cevoli, C., Ragni, L., Fabbri, A., 2021. In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging. Biosystems Engineering 207, 59–67.

Bird, J.J., Barnes, C.M., Manso, L.J., Ekárt, A., Faria, D.R., 2022. Fruit quality and defect image classification with conditional GAN data augmentation. Scientia Horticulturae 293, 110684.

Bramley, R., Pearse, B., Chamberlain, P., 2003. Being Profitable Precisely -A case study of Precision Viticulture from Margaret River. Australian and New Zealand Grapegrower and Winemaker [Annual Technical Issue] 473a, 84–87.

Bramley, R.G., Hamilton, R.P., 2007. Terroir and precision viticulture: are they compatible? OENO One 41, 1–8.

Bramley, R.G.V., 2022. Precision Viticulture: Managing vineyard variability for improved quality outcomes, in: Managing Wine Quality. Elsevier, pp. 541–586.

Bramley, R.G.V., 2005. Understanding variability in winegrape production systems 2. Within vineyard variation in quality over several vintages. Australian Journal of Grape and Wine Research 11, 33–42.

Bramley, R.G.V., Ouzman, J., Boss, P.K., 2011a. 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, 217–229.

Bramley, R.G.V., Ouzman, J., Trought, M.C., Neal, S.M., Bennett, J.S., 2019. Spatio‐temporal variability in vine vigour and yield in a Marlborough Sauvignon Blanc vineyard. Australian Journal of Grape and Wine Research 25, 430–438.

Bramley, R.G.V., Proffitt, A.P.B., Hinze, C.J., Pearse, B., Hamilton, R.P., 2005. Generating benefits from precision viticulture through selective harvesting. Precision agriculture 5, 891–898.

Bramley, R.G.V., Trought, M.C., Praat, J.-P., 2011b. Vineyard variability in Marlborough, New Zealand: characterising variation in vineyard performance and

options for the implementation of Precision Viticulture. Australian Journal of Grape and Wine Research 17, 72–78.

Breiman, L., 1996. Bagging predictors. Machine learning 24, 123–140.

Brook, A., De Micco, V., Battipaglia, G., Erbaggio, A., Ludeno, G., Catapano, I., Bonfante, A., 2020. A smart multiple spatial and temporal resolution system to support precision agriculture from satellite images: Proof of concept on Aglianico vineyard. Remote Sensing of Environment 240, 111679.

Cai, W., Li, Y., Shao, X., 2008. A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra. Chemometrics and intelligent laboratory systems 90, 188–194.

Carrillo, E., Matese, A., Rousseau, J., Tisseyre, B., 2016. Use of multi-spectral airborne imagery to improve yield sampling in viticulture. Precision agriculture 17, 74–92.

Chandrasekaran, I., Panigrahi, S.S., Ravikanth, L., Singh, C.B., 2019. Potential of near-infrared (NIR) spectroscopy and hyperspectral imaging for quality and safety assessment of fruits: An overview. Food Analytical Methods 12, 2438–2458.

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near‐infrared reflectance spectroscopy–principal components regression analyses of soil properties. Soil Science Society of America Journal 65, 480–490.

Chen, D., Qi, X., Zheng, Y., Lu, Y., Huang, Y., Li, Z., 2024. Synthetic data augmentation by diffusion probabilistic models to enhance weed recognition. Computers and Electronics in Agriculture 216, 108517.

Chen, S., Zhang, F., Ning, J., Liu, X., Zhang, Z., Yang, S., 2015. Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging. Food Chemistry 172, 788–793.

Cogato, A., Pagay, V., Marinello, F., Meggio, F., Grace, P., De Antoni Migliorati, M., 2019. Assessing the feasibility of using sentinel-2 imagery to quantify the impact of heatwaves on irrigated vineyards. Remote Sensing 11, 2869.

Coombe, B.G., 1995. Growth stages of the grapevine: adoption of a system for identifying grapevine growth stages. Australian journal of grape and wine research 1, 104–110.

Cozzolino, D., Cynkar, W.U., Shah, N., Smith, P., 2011. Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality. Food Research International 44, 1888–1896.

Dambergs, R., Gishen, M., Cozzolino, D., 2015. A review of the state of the art, limitations, and perspectives of infrared spectroscopy for the analysis of wine grapes, must, and grapevine tissue. Applied Spectroscopy Reviews 50, 261–278.

Dambergs, R.G., Cozzolino, D., Esler, M.B., 2003. The use of near infrared spectroscopy for rape quality measurement. Australian and New Zealand grapegrower and winemaker 69–76.

Danielmeier, T., 2008. The Winescape of New Zealand. Pac. News 29, 11–13.

Das, P., Yadav, J.P.S., 2020. Transfer learning based tomato ripeness classification, in: 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE, pp. 423–428.

Delgadillo-Duran, D.A., Vargas-García, C.A., Varón-Ramírez, V.M., Calderón, F., Montenegro, A.C., Reyes-Herrera, P.H., 2022. Vis–NIR spectroscopy and machine learning methods to diagnose chemical properties in Colombian sugarcane soils. Geoderma Regional 31, e00588.

dos Santos Costa, D., Mesa, N.F.O., Freire, M.S., Ramos, R.P., Mederos, B.J.T., 2019. Development of predictive models for quality and maturation stage attributes of wine grapes using vis-nir reflectance spectroscopy. Postharvest Biology and Technology 150, 166–178.

Eichhorn, K.W., Lorenz, D.H., 1977. Phanologische entwicklungsstadien der rebe. Nachrichtenblatt des Deutschen Pflanzenschutzdienstes.

Elmasry, G., Kamruzzaman, M., Sun, D.-W., Allen, P., 2012. Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. Critical reviews in food science and nutrition 52, 999–1023.

Fairuz Omar, A., 2013. Spectroscopic profiling of soluble solids content and acidity of intact grape, lime, and star fruit. Sensor Review 33, 238–245.

Fernandes, A.M., Franco, C., Mendes-Ferreira, A., Mendes-Faia, A., da Costa, P.L., Melo-Pinto, P., 2015. Brix, pH and anthocyanin content determination in whole Port wine grape berries by hyperspectral imaging and neural networks. Computers and Electronics in Agriculture 115, 88–96.

Ferrer-Gallego, R., Hernández-Hierro, J.M., Rivas-Gonzalo, J.C., Escribano-Bailón, M.T., 2011. Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. LWT-Food Science and Technology 44, 847–853.

Ferro, M.V., Catania, P., Miccichè, D., Pisciotta, A., Vallone, M., Orlando, S., 2023. Assessment of vineyard vigour and yield spatio-temporal variability based on UAV high resolution multispectral images. biosystems engineering 231, 36–56.

Froment, M., Dampney, P., Goodlass, G., Dawson, C., Clarke, J., 1995. A review of spatial variation of nutrients in soil. MAFF final report for project CE0139, ministry of agriculture London, UK: Fisheries and Food.

García-Fernández, M., Sanz-Ablanedo, E., Rodríguez-Pérez, J.R., 2021. High-resolution drone-acquired RGB imagery to estimate spatial grape quality variability. Agronomy 11, 655.

Gil-Pérez, B., Zarco-Tejada, P.J., Correa-Guimaraes, A., Relea-Gangas, E., Navas-Gracia, L.M., Hernández-Navarro, S., Sanz-Requena, J.F., Berjón, A., Martín-Gil, J., 2010. Remote sensing detection of nutrient uptake in vineyards using narrow-band hyperspectral imagery. Vitis 49, 167–173.

Giovos, R., Tassopoulos, D., Kalivas, D., Lougkos, N., Priovolou, A., 2021. Remote sensing vegetation indices in viticulture: A critical review. Agriculture 11, 457.

Golic, M., Walsh, K., Lawson, P., 2003. Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. Applied spectroscopy 57, 139–145.

Gomes, V., Fernandes, A., Faia, A., Melo-Pinto, P., 2017a. Comparison of different approaches for the prediction of sugar content in new vintages of whole Port wine grape berries using hyperspectral imaging. Computers and electronics in agriculture 140, 244–254.

Gomes, V., Fernandes, A., Martins-Lopes, P., Pereira, L., Faia, A.M., Melo-Pinto, P., 2017b. Characterization of neural network generalization in the determination of pH and anthocyanin content of wine grape in new vintages and varieties. Food chemistry 218, 40–46.

Gomes, V., Mendes-Ferreira, A., Melo-Pinto, P., 2021a. Application of hyperspectral imaging and deep learning for robust prediction of sugar and pH levels in wine grape berries. Sensors 21, 3459.

Gomes, V., Rendall, R., Reis, M.S., Mendes-Ferreira, A., Melo-Pinto, P., 2021b. Determination of sugar, pH, and anthocyanin contents in port wine grape berries through hyperspectral imaging: An extensive comparison of linear and non-linear predictive methods. Applied Sciences 11, 10319.

González-Caballero, V., Pérez-Marín, D., López, M.-I., Sánchez, M.-T., 2011. Optimization of NIR spectral data management for quality control of grape bunches during on-vine ripening. Sensors 11, 6109–6124.

González-Caballero, V., Sánchez, M.-T., Fernández-Novales, J., López, M.-I., Pérez-Marín, D., 2012. On-vine monitoring of grape ripening using near-infrared spectroscopy. Food Analytical Methods 5, 1377–1385.

González-Caballero, V., Sánchez, M.-T., López, M.-I., Pérez-Marín, D., 2010. First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery. Journal of Food Engineering 101, 158–165.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. Advances in neural information processing systems 27.

Gour, M., Jain, S., Sunil Kumar, T., 2020. Residual learning based CNN for breast cancer histopathological image classification. International Journal of Imaging Systems and Technology 30, 621–635.

Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G., Frias, J.M., 2007. Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. Trends in food science & technology 18, 590–598.

Grafton, M., Kaul, T., Palmer, A., Bishop, P., White, M., 2019. Regression analysis of proximal hyperspectral data to predict soil pH and Olsen P. Agriculture 9, 55.

Grainger, K., Tattersall, H., 2016. Wine production and quality. John Wiley & Sons.

Greensill, C.V., Walsh, K.B., 2000. A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit. Measurement Science and Technology 11, 1674.

Guidetti, R., Beghi, R., Bodria, L., 2010. Evaluation of grape quality parameters by a simple Vis/NIR system. Transactions of the ASABE 53, 477–484.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C., 2017. Improved training of wasserstein gans. Advances in neural information processing systems 30.

Guo, Z., Zheng, H., Xu, X., Ju, J., Zheng, Z., You, C., Gu, Y., 2021. Quality grading of jujubes using composite convolutional neural networks in combination with RGB color space segmentation and deep convolutional generative adversarial networks. Journal of Food Process Engineering 44, e13620.

Gutiérrez, S., Tardáguila, J., Fernández‐Novales, J., Diago, M.P., 2019. On‐the‐go hyperspectral imaging for the in‐field estimation of grape berry soluble solids and anthocyanin concentration. Australian journal of grape and wine research 25, 127–133.

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

Hall, A., Louis, J., Lamb, D., 2003. Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images. Computers & Geosciences 29, 813–822.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

Iqbal, A., Sun, D.-W., Allen, P., 2013. Prediction of moisture, color and pH in cooked, pre-sliced turkey hams by NIR hyperspectral imaging system. Journal of Food Engineering 117, 42–51.

Janiesch, C., Zschech, P., Heinrich, K., 2021. Machine learning and deep learning. Electronic Markets 31, 685–695.

Jasse, A., Berry, A., Aleixandre-Tudo, J.L., Poblete-Echeverría, C., 2021. Intra-block spatial and temporal variability of plant water status and its effect on grape and wine parameters. Agricultural Water Management 246, 106696.

Ji, Y., Liu, R., Xiao, Y., Cui, Y., Chen, Z., Zong, X., Yang, T., 2023. Faba bean above-ground biomass and bean yield estimation based on consumer-grade unmanned aerial vehicle RGB images and ensemble learning. Precision Agriculture 1–22.

Jiang, G., Grafton, M.C., Pearson, D., Bretherton, M.R., Holmes, A., 2022. A Comparison of Supervised Machine Learning Algorithms for Predicting Subfield Yield Variability of Maize Grain. Journal of the ASABE 65, 287–294.

Jiménez-Brenes, F.M., Lopez-Granados, F., Torres-Sánchez, J., Peña, J.M., Ramírez, P., Castillejo-González, I.L., de Castro, A.I., 2019. Automatic UAV-based detection of Cynodon dactylon for site-specific vineyard management. PloS one 14, e0218132.

Kalopesa, E., Karyotis, K., Tziolas, N., Tsakiridis, N., Samarinas, N., Zalidis, G., 2023. Estimation of Sugar Content in Wine Grapes via In Situ VNIR–SWIR Point Spectroscopy Using Explainable Artificial Intelligence Techniques. Sensors 23, 1065.

Kasimati, A., Espejo-García, B., Darra, N., Fountas, S., 2022. Predicting grape sugar content under quality attributes using normalized difference vegetation index data and automated machine learning. Sensors 22, 3249.

Kasimati, A., Espejo-Garcia, B., Vali, E., Malounas, I., Fountas, S., 2021. Investigating a selection of methods for the prediction of total soluble solids among wine grape quality characteristics using normalized difference vegetation index data from proximal and remote sensing. Frontiers in Plant Science 12, 683078.

Keller, M., 2020. The science of grapevines. Academic press.

Khalifa, N.E., Loey, M., Mirjalili, S., 2022. A comprehensive survey of recent trends in deep learning for digital images augmentation. Artificial Intelligence Review 55, 2351–2377.

Kim, J., Mowat, A., Poole, P., Kasabov, N., 2000. Linear and non-linear pattern recognition models for classification of fruit from visible–near infrared spectra. Chemometrics and intelligent laboratory systems 51, 201–216.

King, P.D., Smart, R.E., McClellan, D.J., 2014. Within‐vineyard variability in vine vegetative growth, yield, and fruit and wine composition of Cabernet Sauvignon in Hawke's Bay, New Zealand. Australian Journal of Grape and Wine Research 20, 234–246.

Kuai, B., Chen, J., Hörtensteiner, S., 2018. The biochemistry and molecular biology of chlorophyll breakdown. Journal of experimental botany 69, 751–767.

Lafontaine, M., Bockaj, Z., Freund, M., Vieth, K.-U., Negara, C., Längle, T., 2015. Non-destructive determination of grape berry sugar concentration using visible/near infrared imaging and possible impact on wine quality. tm-Technisches Messen 82, 633–642.

Lamb, D.W., Weedon, M.M., Bramley, R.G.V., 2004. Using remote sensing to predict grape phenolics and colour at harvest in a Cabernet Sauvignon vineyard: Timing observations against vine phenology and optimising image resolution. Australian journal of grape and wine research 10, 46–54.

Lambert, J., Edwards, A., 2023. Electromagnetic Radiation [WWW Document]. CLIMATE SCIENCE INVESTIGATIONS (CSI). URL https://www.ces.fau.edu/nasa/module-2/radiation-sun.php (accessed 11.6.23).

Larraín, M., Guesalaga, A.R., Agosín, E., 2008. A multipurpose portable instrument for determining ripeness in wine grapes using NIR spectroscopy. IEEE Transactions on Instrumentation and Measurement 57, 294–302.

LeCun, Y., Jackel, L., Bottou, L., Brunot, A., Cortes, C., Denker, J., Drucker, H., Guyon, I., Muller, U.A., Sackinger, E., 1995. Comparison of learning algorithms for handwritten digit recognition, in: International Conference on Artificial Neural Networks. Perth, Australia, pp. 53–60.

Lee, H., Kwon, H., 2017. Going deeper with contextual CNN for hyperspectral image classification. IEEE Transactions on Image Processing 26, 4843–4855.

Lee, J.-H., Kang, T.H., Um, B.H., Sohn, E.-H., Han, W.-C., Ji, S.-H., Jang, K.-H., 2013. Evaluation of physicochemical properties and fermenting qualities of apple wines added with medicinal herbs. Food Science and Biotechnology 22, 1039–1046.

Li, M., 2017. Developing non-destructive techniques to predict 'Hayward' kiwifruit storability. Massey University, New Zealand.

Li, S., Luo, H., Hu, M., Zhang, M., Feng, J., Liu, Y., Dong, Q., Liu, B., 2019. Optical non-destructive techniques for small berry fruits: A review. Artificial Intelligence in Agriculture 2, 85–98.

Lu, Y., Chen, D., Olaniyi, E., Huang, Y., 2022. Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. Computers and Electronics in Agriculture 200, 107208.

Lu, Y., Saeys, W., Kim, M., Peng, Y., Lu, R., 2020. Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. Postharvest Biology and Technology 170, 111318.

Lu, Y., Young, S., 2020. A survey of public datasets for computer vision tasks in precision agriculture. Computers and Electronics in Agriculture 178, 105760.

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Sandoval, E., 2024. Hyperspectral Imaging Spectroscopy for Non-Destructive Determination of Grape Berry Total Soluble Solids and Titratable Acidity. Remote Sensing 16, 1655.

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Sandoval, E., 2023a. Assessing the Leaf Blade Nutrient Status of Pinot Noir Using Hyperspectral Reflectance and Machine Learning Models. Remote Sensing 15, 1497.

Lyu, H., Grafton, M., Ramilan, T., Irwin, M., Wei, H.-E., Sandoval, E., 2023b. Using Remote and Proximal Sensing Data and Vine Vigor Parameters for Non-Destructive and Rapid Prediction of Grape Quality. Remote Sensing 15, 5412.

Manley, M., 2014. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chemical Society Reviews 43, 8200–8214.

Martens, H., Jensen, S.A., Geladi, P., 1983. Multivariate linearity transformation for near-infrared reflectance spectrometry, in: Proceedings of the Nordic Symposium on Applied Statistics. Stokkand Forlag Publishers Stavanger, Norway, pp. 205–234.

Martins, J.A., Guerra, R., Pires, R., Antunes, M.D., Panagopoulos, T., Brázio, A., Afonso, A.M., Silva, L., Lucas, M.R., Cavaco, A.M., 2022. SpectraNet–53: A deep residual learning architecture for predicting soluble solids content with VIS–NIR spectroscopy. Computers and Electronics in Agriculture 197, 106945.

Matese, A., Toscano, P., Di Gennaro, S.F., Genesio, L., Vaccari, F.P., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R., Gioli, B., 2015. Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. Remote Sensing 7, 2971–2990.

Maulud, D., Abdulazeez, A.M., 2020. A review on linear regression comprehensive in machine learning. Journal of Applied Science and Technology Trends 1, 140–147.

Merzlyak, M.N., Solovchenko, A.E., Smagin, A.I., Gitelson, A.A., 2005. Apple flavonols during fruit adaptation to solar radiation: spectral features and technique for non-destructive assessment. Journal of plant physiology 162, 151–160.

Milton, E.J., 2009. Field spectroscopy, in: Geoinformatics. EOLSS Publishers, Oxford, United Kingdom, pp. 209–238.

Milton, E.J., Schaepman, M.E., Anderson, K., Kneubühler, M., Fox, N., 2009. Progress in field spectroscopy. Remote Sensing of Environment 113, S92–S109.

Miranda, J.C., Gené-Mola, J., Zude-Sasse, M., Tsoulias, N., Escolà, A., Arnó, J., Rosell-Polo, J.R., Sanz-Cortiella, R., Martínez-Casasnovas, J.A., Gregorio, E., 2023. Fruit sizing using AI: a review of methods and challenges. Postharvest Biology and Technology 206, 112587.

Moghimi, A., Pourreza, A., Zuniga-Ramirez, G., Williams, L.E., Fidelibus, M.W., 2020. A novel machine learning approach to estimate grapevine leaf nitrogen concentration using aerial multispectral imagery. Remote Sensing 12, 3515.

Moreda, G.P., Ortiz-Cañavate, J., García-Ramos, F.J., Ruiz-Altisent, M., 2009. Non-destructive technologies for fruit and vegetable size determination–a review. Journal of Food Engineering 92, 119–136.

Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R.J., Fredes, C., Valenzuela, A., 2020. A review of convolutional neural network applied to fruit image processing. Applied Sciences 10, 3443.

Ncama, K., Opara, U.L., Tesfay, S.Z., Fawole, O.A., Magwaza, L.S., 2017. Application of Vis/NIR spectroscopy for predicting sweetness and flavour parameters of 'Valencia'orange (Citrus sinensis) and 'Star Ruby'grapefruit (Citrus x paradisi Macfad). Journal of Food Engineering 193, 86–94.

New Zealand Winegrowers, 2023. Vineyard Report 2023 New Zealand Winegrowers. New Zealand Winegrowers, New Zealand.

Newson, D.N., Ratcliff, A.R., Freckleton, J.C., 2012. Practical applications of precision viticulture in Australia, in: I International Workshop on Vineyard Mechanization and Grape and Wine Quality 978. pp. 37–46.

Nicholas, K.A., 2015. Will we still enjoy Pinot noir? Scientific American 312, 60–67.

Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007a. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest biology and technology 46, 99–118.

Nicolai, B.M., Theron, K.I., Lammertyn, J., 2007b. Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple. Chemometrics and intelligent laboratory systems 85, 243–252.

Noriega, L., 2005. Multilayer perceptron tutorial. School of Computing. Staffordshire University 4, 444.

Pádua, L., Marques, P., Hruška, J., Adão, T., Bessa, J., Sousa, A., Peres, E., Morais, R., Sousa, J.J., 2018. Vineyard properties extraction combining UAS-based RGB imagery with elevation data. International Journal of Remote Sensing 39, 5377–5401.

Proffitt, T., Malcolm, A., 2005. Implementing zonal vineyard management through airborne remote sensing. The Australian & New Zealand Grapegrower and Winemaker 22.

Pu, Y.-Y., Feng, Y.-Z., Sun, D.-W., 2015. Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review. Comprehensive Reviews in Food Science and Food Safety 14, 176–188.

Pullanagari, R.R., Dehghan-Shoar, M., Yule, I.J., Bhatia, N., 2021. Field spectroscopy of canopy nitrogen concentration in temperate grasslands using a convolutional neural network. Remote Sensing of Environment 257, 112353.

Pullanagari, R.R., Li, M., 2021. Uncertainty assessment for firmness and total soluble solids of sweet cherries using hyperspectral imaging and multivariate statistics. Journal of Food Engineering 289, 110177.

Ramos, R.P., Gomes, J.S., Prates, R.M., Simas Filho, E.F., Teruel, B.J., dos Santos Costa, D., 2021. Non‐invasive setup for grape maturation classification using deep learning. Journal of the Science of Food and Agriculture 101, 2042–2051.

Rey-Caramés, C., Diago, M.P., Martín, M.P., Lobo, A., Tardaguila, J., 2015. Using RPAS multi-spectral imagery to characterise vigour, leaf development, yield components and berry composition variability within a vineyard. Remote sensing 7, 14458–14481.

Reynolds, A.G., Senchuk, I.V., van der Reest, C., De Savigny, C., 2007. Use of GPS and GIS for elucidation of the basis for terroir: Spatial variation in an Ontario Riesling vineyard. American Journal of Enology and Viticulture 58, 145–162.

Rinnan, Å., Van Den Berg, F., Engelsen, S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends in Analytical Chemistry 28, 1201–1222.

Rolle, L., Segade, S.R., Paissoni, M.A., Giacosa, S., Gerbi, V., 2022. Assessment and control of grape maturity and quality, in: White Wine Technology. Elsevier, pp. 1–16.

Romero, M., Luo, Y., Su, B., Fuentes, S., 2018. Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. Computers and electronics in agriculture 147, 109–117.

Rouse Jr, J.W., Haas, R.H., Deering, D.W., Schell, J.A., Harlan, J.C., 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.

Sanches, I.D., Souza Filho, C.R., Kokaly, R.F., 2014. Spectroscopic remote sensing of plant stress at leaf and canopy levels using the chlorophyll 680 nm

absorption feature with continuum removal. ISPRS Journal of Photogrammetry and Remote Sensing 97, 111–122.

Shaw, T.B., 2012. A climatic analysis of wine regions growing pinot noir. Journal of wine research 23, 203–228.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of big data 6, 1–48.

Silva, R., Gomes, V., Mendes-Faia, A., Melo-Pinto, P., 2018. Using support vector regression and hyperspectral imaging for the prediction of oenological parameters on different vintages and varieties of wine grape berries. Remote Sensing 10, 312.

Silva, R., Melo-Pinto, P., 2023. t-SNE: A study on reducing the dimensionality of hyperspectral data for the regression problem of estimating oenological parameters. Artificial Intelligence in Agriculture 7, 58–68.

Silva, R., Melo-Pinto, P., 2021. A review of different dimensionality reduction methods for the prediction of sugar content from hyperspectral images of wine grape berries. Applied Soft Computing 113, 107889.

Soubry, I., Patias, P., Tsioukas, V., 2017. Monitoring vineyards with UAV and multi-sensors for the assessment of water stress and grape maturity. Journal of Unmanned Vehicle Systems 5, 37–50.

SU, S.L., Singh, D.N., Baghini, M.S., 2014. A critical review of soil moisture measurement. Measurement 54, 92–105.

Sun, D.-W., 2010. Hyperspectral imaging for food quality analysis and control. Elsevier.

Sun, L., Gao, F., Anderson, M.C., Kustas, W.P., Alsina, M.M., Sanchez, L., Sams, B., McKee, L., Dulaney, W., White, W.A., 2017. Daily mapping of 30 m LAI and NDVI for grape yield prediction in California vineyards. Remote Sensing 9, 317.

Tan, H., Hu, Y., Ma, B., Yu, G., Li, Y., 2024. An improved DCGAN model: Data augmentation of hyperspectral image for identification pesticide residues of Hami melon. Food Control 157, 110168.

Taylor, J., Tisseyre, B., Bramley, R., Reid, A., 2005. A comparison of the spatial variability of vineyard yield in European and Australian production systems, in: Precision Agriculture ‘05. Wageningen Academic, pp. 907–914.

Tisseyre, B., Mazzoni, C., Fonta, H., 2008. Whithin-field temporal stability of some parameters in viticulture: Potential toward a site specific management. OENO One 42, 27–39.

Trought, M.C., Bramley, R.G., 2011. Vineyard variability in Marlborough, New Zealand: characterising spatial and temporal changes in fruit composition and juice quality in the vineyard. Australian Journal of Grape and Wine Research 17, 79–89.

Trought, M.C., Dixon, R., Mills, T., Greven, M., Agnew, R., Mauk, J.L., Praat, J.-P., 2008. The impact of differences in soil texture within a vineyard on vine vigour, vine earliness and juice composition. OENO One 42, 67–72.

Tsakiridis, N.L., Samarinas, N., Kokkas, S., Kalopesa, E., Tziolas, N.V., Zalidis, G.C., 2023. In situ grape ripeness estimation via hyperspectral imaging and deep autoencoders. Computers and Electronics in Agriculture 212, 108098.

Urraca, R., Sanz‐Garcia, A., Tardaguila, J., Diago, M.P., 2016. Estimation of total soluble solids in grape berries using a hand‐held NIR spectrometer under field conditions. Journal of the Science of Food and Agriculture 96, 3007–3016.

van Leeuwen, C., Tregoat, O., Choné, X., Gaudillère, J.P., Pernet, D., 2008. Different environmental conditions, different results: the role of controlled environmental stress on grape quality potential and the way to monitor it, in: Proceedings of the Thirteenth Australian Wine Industry Technical Conference; Blair, R., Williams, P., Pretorius, S., Eds. pp. 39–46.

Walsh, K.B., Blasco, J., Zude-Sasse, M., Sun, X., 2020. Visible-NIR 'point'spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. Postharvest Biology and Technology 168, 111246.

Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., Li, X., 2021. A review of deep learning used in the hyperspectral image analysis for agriculture. Artificial Intelligence Review 54, 5205–5253.

Wang, F., Zhao, C., Yang, H., Jiang, H., Li, L., Yang, G., 2022. Non-destructive and in-site estimation of apple quality and maturity by hyperspectral imaging. Computers and Electronics in Agriculture 195, 106843.

Webster, R., Oliver, M.A., 2007. Geostatistics for environmental scientists. John Wiley & Sons.

Wei, H.-E., Grafton, M., Bretherton, M., Irwin, M., Sandoval, E., 2023. Evaluation of the use of two-stage calibrated PlanetScope images and environmental variables for the development of the grapevine water status prediction model. Technology in Agronomy 3.

Wei, H.-E., Grafton, M., Bretherton, M., Irwin, M., Sandoval, E., 2022. Evaluation of the use of UAV-derived vegetation indices and environmental variables for grapevine water status monitoring based on machine learning algorithms and SHAP analysis. Remote Sensing 14, 5918.

Wei, H.-E., Grafton, M., Bretherton, M., Irwin, M., Sandoval, E., 2021. Evaluation of Point Hyperspectral Reflectance and Multivariate Regression Models for Grapevine Water Status Estimation. Remote Sensing 13, 3198. https://doi.org/10.3390/rs13163198

Williams, P., Norris, K., 1987. Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists, Inc.

Willwerth, J.J., Reynolds, A.G., 2020. Spatial variability in Ontario Riesling Vineyards: I. Soil, vine water status and vine performance. Oeno One 54, 327–349.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems 58, 109–130.

Wolf, T.K., 2008. Wine Grape Production Guide for Eastern North America. NRAES-145. Cooperative Extension, Ithaca, NY.

Wolpert, D.H., 1992. Stacked generalization. Neural networks 5, 241–259.

Wu, D., Sun, D.-W., 2013. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part I: Fundamentals. Innovative Food Science & Emerging Technologies 19, 1–14.

Xiaobo, Z., Jiewen, Z., Povey, M.J., Holmes, M., Hanpin, M., 2010. Variables selection methods in near-infrared spectroscopy. Analytica chimica acta 667, 14–32.

Yinshan, G., Zaozhu, N., Kai, S., Jia, Z., Zhihua, R., Yuhui, Z., Quan, G., Hongyan, G., Xiuwu, G., 2017. Composition and content analysis of sugars and organic acids for 45 grape cultivars from northeast region of China. Pak. J. Bot 49, 155–160.

Yu, R., Brillante, L., Martínez-Lüscher, J., Kurtural, S.K., 2020. Spatial variability of soil and plant water status and their cascading effects on grapevine physiology are linked to berry and wine chemistry. Frontiers in Plant Science 11, 790.

Zhang, A., Lipton, Z.C., Li, M., Smola, A.J., 2023. Dive into deep learning. Cambridge University Press.

Zhang, X., Lin, T., Xu, J., Luo, X., Ying, Y., 2019. DeepSpectra: An end-to-end deep learning approach for quantitative spectral analysis. Analytica chimica acta 1058, 48–57.

Zhang, X., Yang, J., Lin, T., Ying, Y., 2021. Food and agro-product quality evaluation based on spectroscopy and deep learning: A review. Trends in Food Science & Technology 112, 431–441.

Zhao, M., Cang, H., Chen, H., Zhang, C., Yan, T., Zhang, Y., Gao, P., Xu, W., 2023. Determination of quality and maturity of processing tomatoes using near-infrared hyperspectral imaging with interpretable machine learning methods. LWT 183, 114861.

Zhao, X., Zhao, D., Wang, J., Triantafilis, J., 2022. Soil organic carbon (SOC) prediction in Australian sugarcane fields using Vis–NIR spectroscopy with different model setting approaches. Geoderma Regional 30, e00566.

Zhao, Z., Hicks, Y., Sun, X., Luo, C., 2023. Peach ripeness classification based on a new one-stage instance segmentation model. Computers and Electronics in Agriculture 214, 108369.

Zhong, Z., Li, J., Luo, Z., Chapman, M., 2017. Spectral–spatial residual network for hyperspectral image classification: A 3-D deep learning framework. IEEE Transactions on Geoscience and Remote Sensing 56, 847–858.

9 Appendix

9.1 Appendix 1

Table A1: Hyperparameters setting of different machine learning models

Model	Hyperparameter
Lasso regression	alpha = 1 shrinkage factor = 0:0.5 nlambda = 1:100 lambda.min.ratio = 0:0.05
Ridge regression	alpha = 0 shrinkage factor = 0:0.5 nlambda = 1:100 lambda.min.ratio = 0:0.05
KNN	k = 3:30 distance = “L1”, “L2”, “generalized distance metric” kernel = "rectangular", "gaussian","rank", "optimal"
SVR	cost = 2^{-2}:2^{1} gamma = 0.01:0.04 epsilon = 0.01:0.5
RFR	num.trees = 250, 500, 750, 1000 min.node.size = 1:10 mtry = 1:10
XGBoost	eta = 0.01:0.2 subsample = 0.5:0.8 colsample_bylevel = 0.5 ;0.7 nrounds = 50:300 max_depth = 2:8 lambda = 0:0.5 alpha = 0:0.005
ANN	nnet.size = 5:30 nnet. decay = 0.1:0.5 nnet. MaxNWts = 2000:3000

9.2 Appendix 2

Table A2: Hyperparameters setting of different machine learning models

Regression model	Hyperparameter	Range
KNN	Number of PCs	1-20
	Number of Nearest Neighbors	log3-log30
	Kernel Type	"rectangular", "gaussian", "rank", "optimal"
	Distance Metric	1-3
RF	Number of PCs	1-20
	Number of Features per Split	2-20
	Number of Trees	1-4
	Minimum Node Size	1-10
SVR	Number of PCs	1-20
	Cost Parameter	2^{-2-7}
	Gamma Parameter	0-0.005
ES	Number of PCs	1-20
	Number of Nearest Neighbors	log3-log30
	Kernel Type	"gaussian", "optimal"
	Distance Metric	1-3
	Number of Features per Split	2-20
	Number of Trees	1-4
	Minimum Node Size	1-10
	Cost Parameter	2^{-2-5}
	Gamma Parameter	0-0.003

9.3 Appendix 3

Table A3: Hyperparameter setting of 3DCNN

Hyperparameter		Value
Convolutional layer 1	Input channels	1
	Output channels	32
	Kernel size	4,4,32
	Padding	3,3,0
Convolutional layer 2	Input channels	32
	Output channels	32
	Kernel size	5,5,32
Convolutional layer 3	Input channels	32
	Output channels	32
	Kernel size	4,4,32
	Padding	2,2,0
Max pooling	Kernel size	2,2,0
Dropout		0.5
Learning rate		0.00001
Epoch		100
Batch size		512
Loss function		CrossEntropyLoss

Optimizer		Adam

Table A4: Hyperparameter setting of SSRN

Hyperparameter		Value
Convolutional layer 1	Input channels	1
	Output channels	24
	Kernel size	1,1,7
	Stride	1,1,2
BatchNorm3D+ReLU	eps	0.001
	momentum	0.1
Residual block 1	Input channels	24
	Output channels	24
	Kernel size	1,1,7
	Padding	0,0,3
Residual block 2	Input channels	24
	Output channels	24
	Kernel size	1,1,7
	Padding	0,0,3
Convolutional layer 2	Input channels	24
	Output channels	128

	Kernel size	1,1,90
	Stride	1,1,1
BatchNorm3D+ReLU	eps	0.001
	momentum	0.1
Convolutional layer 3	Input channels	1
	Output channels	24
	Kernel size	3,3,128
	Stride	1,1,1
BatchNorm3D+ReLU	eps	0.001
	momentum	0.1
Residual block 3	Input channels	24
	Output channels	24
	Kernel size	3,3,1
	Padding	1,1,0
Residual block 4	Input channels	24
	Output channels	24
	Kernel size	3,3,1
	Padding	1,1,0
Average pooling	Kernel size	5,5,1
Learning rate		0.00001

Epoch		100
Batch size		512
Loss function		CrossEntropyLoss
Optimizer		Adam