# Estimating Grapevine Water Status Using Hyperspectral and Multispectral Remote Sensing and Machine Learning Algorithms

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.

Hsiang-En Wei

2023

## Abstract

Moderate water deficit is desirable for achieving the optimal grape composition which determines the values of wine, especially for red cultivars. To attain consistency in grape quality in vineyards, it is critical to manage grapevine water status (GWS) to the target range, but to avoid severe dehydration, between fruit set and veraison. Together with the foreseeable climate changes and stricter environmental regulations, there is a need for viticulture to estimate GWS variability across fields along growing seasons before irrigating to eliminate the uncertainty of controlling hydration status and to produce consistent grapes with premium quality.

Precision viticulture (PV) recognizes that not all areas within a vineyard are uniform in terms of their soil, climate, and other environmental conditions. Therefore, it tailors viticultural management to the unique needs of different vineyard zones by focusing on applying site-specific or time-specific management practices. PV aims to enhance grape quality and yield while minimizing resource usage and environmental impacts. As the final product (wine) for viticulture has the potential of high additional values, it is worth considering the application of PV in decision-making according to the information on spatio-temporal variability across the fields.

The advancement and availability of remotely sensed spectral information, geospatial technologies, and machine learning models have opened a new chapter for spatio-temporal GWS monitoring. However, there are technical shortcomings that need to be addressed before extensive application and adoption of these techniques in viticulture. These include a lack of understanding of GWS-related spectral data analysis methods, a lack of data interoperability for GWS estimation between data sourced from various devices with different formats, and a lack of availability of high-coverage images with high spatial and temporal resolution. Therefore, this study tackled the technical bottlenecks, related to the application of proximal sensing and remote sensing (RS) in GWS estimation, from three perspectives: (i) the exploration of relevant spectral regions over the electromagnetic spectrum, (ii) the complementation from differently sourced datasets other than RS information, (iii) the provision of large-scale GWS prediction.

This study was undertaken in two Pinot Noir vineyards trained with vertical shoot positioning for two growing seasons in Martinborough, New Zealand. The investigation window, corresponding to the critical periods of GWS management, between fruit set and veraison in each growing season, includes November, December, January, and February. Stem water potential ($\Psi_{stem}$), serving as a proxy for GWS, is measured on 85 and 63 canopies in the first and second growing seasons. Each sampled grapevine is recorded for its location with a global navigation satellite system with real-time kinematic correction. Five times field data collection, including measuring hyperspectral point data using ASD FieldSpec 4 Spectroradiometer (proximal sensing data) and taking multispectral images using DJI Phantom 4 UAV (remote sensing data), were carried out in each growing season. An electromagnetic induction survey was implemented by using EM38-MK2 to acquire apparent electrical conductivity ($EC_a$) maps (complementary data). Several satellite images collected by PlanetScope (remote sensing data) during the study periods and the LiDAR-based digital elevation model (complementary data) were downloaded and added to the analysis datasets. An on-site weather station continuously records and provides meteorological

information, including air temperature (°C), relative humidity (%), rainfall (mm), wind speed (km/h), and irradiance ($W/m^2$) (complementary data).

The identification of the relationships between spectral information and $\Psi_{stem}$ is an essential step for the robust application. By analyzing hyperspectral spectra, it shows that the statistically relevant wavelengths disperse across visible, near-infrared, and shortwave infrared (SWIR) spectral bands. They are specifically located around blue, red, and red edge bands, two weak water absorption bands at 970 and 1200 nm, two strong absorption bands at 1400 and 1940 nm, and some dry matter-related bands. When analyzing multispectral images taken by UAV, it shows Transformed Chlorophyll Absorption Reflectance Index and Excess Green Index are the multispectral indices mostly correlated ($R^2$ = 0.35 and 0.3, respectively) with the changes in $\Psi_{stem}$. It implies that the variation in leaf pigments, especially chlorophylls, is better for describing the $\Psi_{stem}$ variation of Pinot Noir than the alteration in canopy structure.

When applying the $\Psi_{stem}$-sensitive spectral bands through airborne or spaceborne platforms, the missing SWIR for most commercial multispectral sensors and the presence of vapor in the air obstruct the usefulness of the $\Psi_{stem}$-sensitive spectral bands. Therefore, this study assesses the complementary effects provided by other environmental aspects, including soil/ terrain, vegetation, temporal, and weather variables, to improve the GWS estimating capabilities of aerial multispectral sensors. The results prove the complementary effects by displaying that the detection accuracy is improved from RMSE of 213 to 146 kPa and RMSE of 221 kPa to 138 kPa.

To monitor the fields at a large-scale using multispectral satellites, it is common to encounter several technical issues: coarse resolution pixels that contain background information, weather dependence, and delay in the image delivery. This study addresses the limitation of coarse spatial resolution using two-stage calibration to scale information provided by ground measurement up to satellite images, along with removing interference from the inter-row components. It demonstrates that satellite images can approximate the collected $\Psi_{stem}$ with high accuracy (RMSE = 59 kPa). To deal with the contamination by weather and delivering delay of image products, a prediction model is established based on the calibrated satellite images and various environmental variables (day of the year, rainfall, potential evapotranspiration, irrigation, fertigation, plucking, trimming, normalized difference vegetation index, $EC_a$, elevation, and slope). The developed model is able to predict $\Psi_{stem}$ trend in an independent growing season with high consistency when compared with the reference (r = 0.89 and 0.87 for the two vineyards, respectively).

## Acknowledgments

When I step on the finish line of my PhD journey, I still cannot believe I have made it so far. The covid-19 pandemic broke out one month after I arrived in NZ, and it soon became a national lockdown for more than two years. As an international student, I cannot leave the country otherwise I would not be allowed to enter NZ again. These rules prevented international travelers intending to visit NZ, including my beloved wife, Jiaqi. The 840-day separation from my wife was a hard time affecting my psychological status, in addition to the ongoing stress brought by the PhD work in the meantime. However, I am not alone in this journey. The encouragement and support I received over the past three years are undoubtedly the main pillars that enable me to proceed and attain this stage.

My wife sits on the top of whom I would like to express my gratitude. She is not only the most helpful partner but also my soul mate who kept encouraging and accompanying me over social media during our separation, until the present after our reunion. Without her company, it would have been impossible for me to go through the entire PhD process. Following my wife, my parents, younger sister, and parents-in-law are my irreplaceable and precious family who consistently serve as a cheer squad to fill me with confidence and motivation whenever there is a need.

My supervisors, Miles Grafton, Mike Bretherton, and Matt Irwin are always there when I need help. Miles is good at sweeping over administrative obstacles while wisely arranging the available funding. Thanks to his effort, it allows me to completely focus on my research and produce high-quality papers. Mike's brilliant proofreading enables me to smoothly hurdle the barriers in publication, and I always learn how to structure sentences and avoid ambiguity after receiving his edits. For the fieldwork, Matt uses his rich in-situ experience to transform the complex methods into understandable procedures over the two study periods, while assuring the positional accuracy and consistency of the collected data. In addition, this project cannot access UAV images without Eduardo Sandoval's professional UAV operation and his great support.

Last by not least, thanks to all the friends I met in Palmerston North. The splendid memories we enjoyed food around the dining table, we picnicked in the park, we did pick-your-own, we played bowling games, and celebrating will always sit a place in my mind. Looking back again, I appreciate everyone I met during these three years. Without you, this PhD journey would never have happened.

# Table of Contents

## List of Figures

## List of Tables

# List of Appendices

# List of Abbreviations and Acronyms

| Abbreviation/Acronym | Full name |
|---|---|
| ABA | Abscisic acid |
| $EC_a$ | Apparent electrical conductivity |
| $r_b$ | Boundary layer resistance |
| $R^2$ | Coefficient of determination |
| CR | Continuum removal |
| CV | Cross-validation |
| DOY | Day of the year |
| DI | Deficit irrigation |
| EN | Elastic Net |
| EMI | Electromagnetic induction |
| ET | Evapotranspiration |
| 1D | First derivative |
| GNSS | Global navigation satellite system |
| GWS | Grapevine water status |
| GIS | Graphical information system |
| IPCC | International Panel on Climate Change |
| kPa | Kilopascal |
| LAI | Leaf area index |
| $\Psi_{leaf}$ | Midday leaf water potential |
| ML | Machine learning |
| mha | Million hectare |
| mhl | Million hectoliter |
| MLP | Multilayer perceptron |
| NIR | Near infrared |
| NDSI | Normalized difference spectral index |
| PLSR | Partial least squares regression |
| PAWC | Plant available water capacity |
| PS | PlanetScope |
| PET | Potential evapotranspiration |
| PV | Precision viticulture |
| $\Psi_{pd}$ | Predawn leaf water potential |
| RFR | Random forest regression |
| RPIQ | Ratio of performance to interquartile range |
| RFECV | Recursive feature elimination based on cross-validation |
| RCP | Representative concentration pathway |
| RS | Remote sensing |
| RTK | Real-time kinematic |
| RMSE | Root mean square error |
| 2D | Second derivative |
| SHAP | Shapley Additive exPlanations |
| SWIR | Shortwave infrared |
| SRI | Simple ratio spectral index |
| $\Psi_{stem}$ | Stem water potential |
| gs | Stomatal conductance |

| | |
|---|---|
| $r_s$ | Stomatal resistance |
| SVR | Support vector regression |
| UAV | Unmanned aerial vehicle |
| VPD | Vapor pressure deficit |
| VIP | Variable importance of projection |
| VI | Vegetation index |
| VIS | Visible |
| VNIR | Visible and near-infrared |

# Chapter 1
# Introduction

## 1.1 Background Concept and Research Domain

### *1.1.1 Precision Agriculture and Precision Viticulture*

Precision agriculture (PA), a modern farming concept by using digital technologies to optimize agricultural production processes, by site and time-specific application of cultivation inputs, is considered the future of farming (Langemeier & Boehlje, 2021; Oliver et al., 2013). PA does not treat every field as homogenous. Instead, it takes into account the unique, site-specific requirements by considering the spatial and temporal variations that exist across fields. The goal of PA is to increase the use efficiency of agricultural inputs while decreasing the impacts on the environment. PA, if properly adopted, may lead to a combination of better quality produce, cost saving, and increased profit and environmental benefits (Onibonoje & Nwulu, 2021). PA is being adopted into expanding practices in fields across the world, with accelerated uptake in the last ten years (Singh et al., 2022).

Precision viticulture (PV) refers to the concept of PA being applied specifically to grape production and development (Proffitt et al., 2006). Vineyard management and winemaking are commonly linked, as the derived grape composition determines the basis for wine properties. Vineyards are commonly characterized as having high heterogeneity of growing conditions due to different physiological responses of vines to soil physio-chemical features (influencing water holding capacity, infiltration rate, and nutrient availability/ uptake), genotypes, training systems, cultivation practices, and weather dynamics (Bramley & Lamb, 2003). This results in variability in vegetative growth across vineyards, subsequently leading to variation in yield and grape quality. A key aspect of PV is understanding the patterns and extent of within-vineyard heterogeneities, and how these variations make consequences on vine growth and grape quality as the basis for customized viticultural management (Figure 1.1). The rationale behind PV is providing site-specific or time-specific practices to satisfy the real needs of vines, in order to obtain desired or expected outcomes meeting market objectives (Bramley, 2022; Proffitt et al., 2006). This is particularly important for viticulture since its final product (wine) with a potential of high added values can be achieved using timely decision-making, based on knowledge of spatial variability in vineyards (Bramley et al., 2005; King et al., 2014). To date, PV has been used in several aspects of viticultural management, including irrigating (Romero et al., 2018), fertilizing (Balafoutis et al., 2017), canopy management (Sassu et al., 2021), and disease detection (Albetis et al., 2017). Research areas in PV mainly focus on four fields as follows (Arnó Satorra et al., 2009; Tisseyre & Taylor, 2005):

-quantification and evaluation of intra-vineyard variability
-delimitation of management zones at the block level based on the analysis and interpretation of this variability
-development of variable rate technologies
-evaluation of opportunities for site-specific vineyard management

This study falls in the first research area, focusing on assessing the application of various PA technologies in monitoring spatial and temporal variability of grapevine water status (GWS) in vineyards.

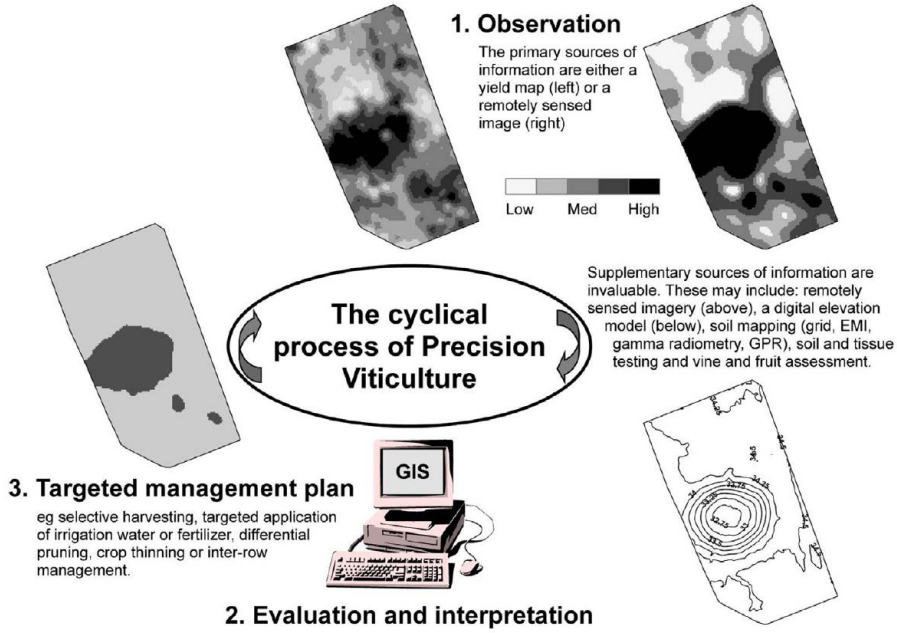



Figure 1. 1 The process of precision viticulture.
Source: Bramley, 2022

### *1.1.2 Viticulture in the World and New Zealand*

The global vineyard planted area for all purposes (non-productive young vines, wine, juice, table grape, and dry grape), is estimated to be approximately 7.3 million hectares (mha) in 2021 (OIV, 2022). In terms of world vineyard planting area, Spain, France, and China are the top three in descending order and account for 964 kha, 798 kha, and 783 kha, respectively. In 2021, wine production, excluding juices and musts, is estimated at 260 million hectoliters (mhL), which is slightly below the past ten-year average. From the perspective of wine consumption, it is estimated at 236 mhl in 2021. The global wine export amount is estimated to be 111.6 mhL in 2021, with an export value of 34.3 billion euros.

In New Zealand (NZ), wine-producing vineyards are distributed across the two main islands, with a total production area of 41,603 ha in 2021 (Winegrowers, 2022). Three cultivars (Sauvignon Blanc, Pinot Noir, and Chardonnay) dominate the vineyards in terms of production area. The most important wine-producing region is Marlborough at the northern end of the South Island, representing more than 70% of the national production area. It also produces over 50% of internationally recognized Sauvignon Blanc, from NZ (Figure 1.2). In 2021, the value of NZ wine exports was nearly 1.8 billion, which makes it the second most important horticultural export in NZ contributing to 27 % of horticultural produce export by value (Plant & Food Research, 2021).

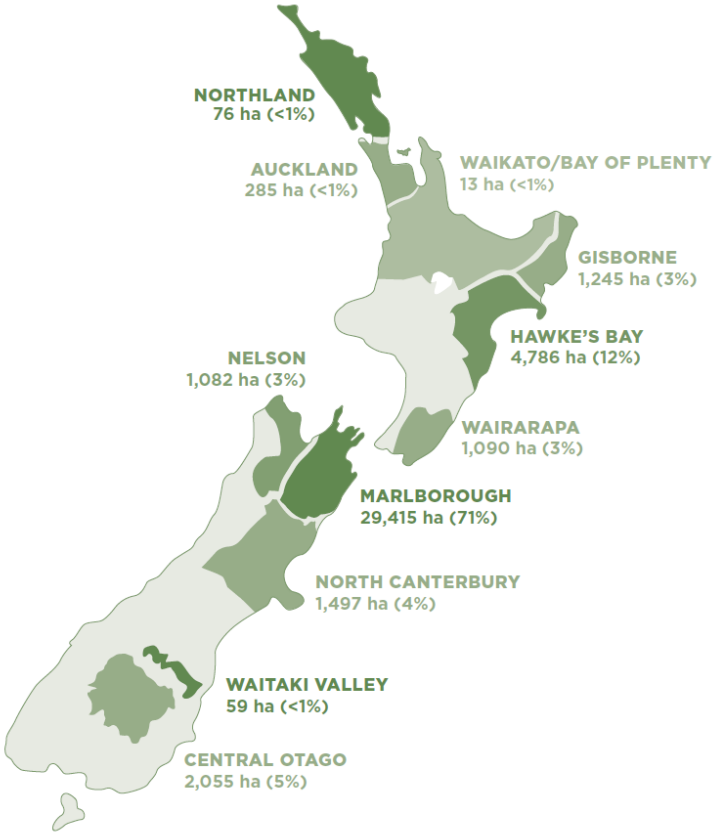



Figure 1. 2 The major vineyard regions, with producing area, of New Zealand.
Source: Winegrowers, 2022

### *1.1.3 The Impacts of Climate Change on Viticulture and Future Climate Scenario in New Zealand*

It is clear that climatic conditions play a dominant role in nearly all agricultural industries, by determining whether a crop can be grown in a given location, and influencing crop productivity and quality. However, climate change has posed a foreseeable challenge to agricultural production, further leading to uncertain outputs and economics. According to the fifth assessment of the International Panel on Climate Change (IPCC), the mean surface temperature is projected to increase at approximately 0.2°C per decade and is expected to rise to values of 1°C to 5°C, above the current temperature, at the end of 21st century (IPCC, 2014). Other climate consequences are likely to be, changes in the precipitation patterns, increasing UV-B radiation, and increasing frequency of occurrence of extreme weather events such as droughts and heatwaves (Solomon et al., 2007; Stocker, 2014).

Grapes are especially sensitive to relatively small variations in climate as perennial vines are planted in one location for at least 30 years, with selected varieties and corresponding management practices that are adapted to the range of local climate variability (Jones, 2006; Jones, 2018). For example, Pinot Noir, as a cool climate variety, is best grown in regions with an average temperature of 14-16 °C. Although Pinot Noir can be grown outside this temperature range, it will lead to grape quality changes that are unknown (Jones, 2021). Given the importance of atmospheric factors to viticulture, it is highly likely that continued climate change would have significant impacts on

vegetative growth, grape yield, quality, and later wine characteristics, which together influence economic sustainability (Fraga & Santos, 2017; Van Leeuwen & Darriet, 2016).

In terms of influencing grapevine water status, increasing temperature is expected to exacerbate water deficit conditions through increasing vapor pressure deficit (VPD) of the air. VPD is the difference between the actual amount of moisture in the air and the amount of moisture in saturated air, and it determines the rate of evapotranspiration (ET) which includes all processes of water movement from land surface to the atmosphere via soil moisture evaporation and plant transpiration. The increasing VPD elevates the evaporative demands on grapevine canopies, which results in greater water requirements from grapevines and a greater negative water balance in the soil (Van Leeuwen et al., 2019). Another cause for increasing evaporative demands will be a greater frequency of heatwave events. A temperature higher than 22 °C to 24 °C is considered to invoke heat stress for vines, which leads to water stress, especially in dry conditions (Santos et al., 2020). Rainfall plays a critical role in controlling soil moisture availability for crops and then determining corresponding GWS, so it is associated with vine phenology, desired yield, and quality of grapes (Neethling et al., 2017). Modifications in rainfall patterns have resulted in changes, such as periods of increasing intensity and less frequent events resulting in extremely dry periods in several regions (Van Leeuwen et al., 2019). Although plants can mitigate the influences of transpiration through the control of the stomatal pores, there is a diversity in stomatal responses to water stress, among different grape genotypes (Levin et al., 2020). These factors provide an incentive for increasing water use efficiency in the viticultural sector, due to increasing difficulties in keeping the more dynamic GWS in an optimal range.

The fifth IPCC assessment developed four scenarios, known as representative concentration pathways (RCPs), including RCP 2.6, 4.5, 6.0, and 8.5 (for respective radiative forcing of 2.6, 4.5, 6.0, and 8.5 W*m-2) (IPCC, 2014). RCPs represent various concentrations of greenhouse gases in the atmosphere, and they are identified by the accumulated radiative forcing at 2100 relative to 1750 based on how much greenhouse gas humans continue to emit. These four pathways are detailed as follows:
-RCP2.6: low concentrations - emissions peak around 2020 and decline substantially thereafter. It is a mitigation pathway.
-RCP4.5: low-mid concentrations - emissions peak around 2040, then declining. It is a stabilization pathway.
-RCP6.0: mid-high concentrations - emissions peak around 2080, then declining. It is a stabilization pathway.
-RCP8.5: very high greenhouse gas concentrations - emissions continue to rise throughout the 21st century, which assumes businesses are operated as usual.

These pathways were downscaled to NZ and presented in Table 1.1 to express NZ climate change projections (MfE, 2018). This table listed the direction and magnitude of the atmospheric variables which have impacts on GWS, in respect of the four RCPs. All the changes are relative to the 1986-2005 period which serves as the baseline historical climate. The focused future periods are labeled as “2040” (the 2031-2050 average) and “2090” (the 2081-2100 average).

Table 1. 1 New Zealand climate change projections for 2040 and 2090.

| Climate variable | Direction of change | Magnitude of change |
|---|---|---|
| Mean temperature | Progressive increase with greenhouse gas concentration. | By 2040, from +0.7°C [RCP2.6] to +1.0°C [RCP8.5]. By 2090, +0.7°C [RCP2.6]to +3.0°C [RCP8.5]. |
| Minimum and maximum temperatures | As mean temperature | Maximum increases faster than the minimum. Diurnal range increases by up to 2°C by 2090 [RCP8.5]. |
| Daily temperature extremes: hot days | Increase in hot days (maximum temperature of 25°C or higher). | By 2040, a 40% [RCP 2.6] to 100% [RCP 8.5] increase. By 2090, a 40% [RCP 2.6] to 300% [RCP 8.5] increase. |
| Mean precipitation | Varies around the country and with season. Annual pattern of increases in the west and south of New Zealand, and decreases in the north and east. | Substantial variation around the country, increasing in magnitude with increasing emissions. |
| Daily precipitation extremes: dry days | More dry days throughout North Island, and in inland South Island. | By 2090 [RCP 8.5], up to 10 or more dry days per year (~5% increase). |
| Drought | Increase in severity and frequency. | By 2090 [RCP 8.5], up to 50mm or more increase per year on average. |
| Extreme wind speeds | Increase | Up to 10% or more in parts of the country |

| Climate variable | Direction of change | Magnitude of change |
|---|---|---|
| Mean temperature | Progressive increase with greenhouse gas concentration. | By 2040, from +0.7°C [RCP2.6] to +1.0°C [RCP8.5]. By 2090, +0.7°C [RCP2.6]to +3.0°C [RCP8.5]. |
| Solar radiation | Varies around the country and with season | Seasonal changes generally lie between -5% and +5% |
| Relative humidity | Decrease | Up to 5% or more by 2090 [RCP 8.5], especially in the South Island |

Source: MfE, 2018

### *1.1.4 Restriction on Water Use*

Vine growth and fruit development are highly dependent on water supply (Jackson & Lombard, 1993). Grapevines have been historically rainfed in the "Old World" (including Europe and the Middle East), and irrigation was prohibited by law in some countries. For instance, irrigation is restricted to the period, from flowering to 15th August (close to the date of veraison), in France's Appellation d'Origine Protégée system to prevent dilution effects on quality-determining compounds (Scholasch & Rienth, 2019). Whilst irrigated vineyards are mostly in the "New World" (referring to regions that are outside the Old World). However, irrigation restrictions have been challenged by global warming, and several countries, including Spain and Italy, have lifted the restrictions to allow irrigation. It has been observed that many vineyards, including some regions under cool climates, have encountered certain degrees of water deficit recently during the growing season, with increasing use of drip irrigation as a result (Van Leeuwen & Seguin, 2006).

NZ amended the Resource Management Regulations for freshwater management and reporting of water uptakes in 2020. Water consent holders are required to increase the frequency of water take reporting (up to a 15-minute period), automatically transferring water take measurements to authorities daily. This generates a more complete database of water use (MfE, 2021). In addition, it introduced increasing minimum flow rates of rivers and capping of water allocation (Figure 1.3). This increased the length of duration when water cannot be harvested, for irrigation from river systems during seasonal low flow (MPI, 2021). Low river flows often corresponds to the peak of water usage and irrigation requirements by agricultural sectors. This alteration imposed substantial impacts on food and fiber production in NZ, as it is the biggest user of freshwater in NZ accounting for 75% of water withdrawal consents, and a high proportion of the irrigation is dependent on run-of-river water.

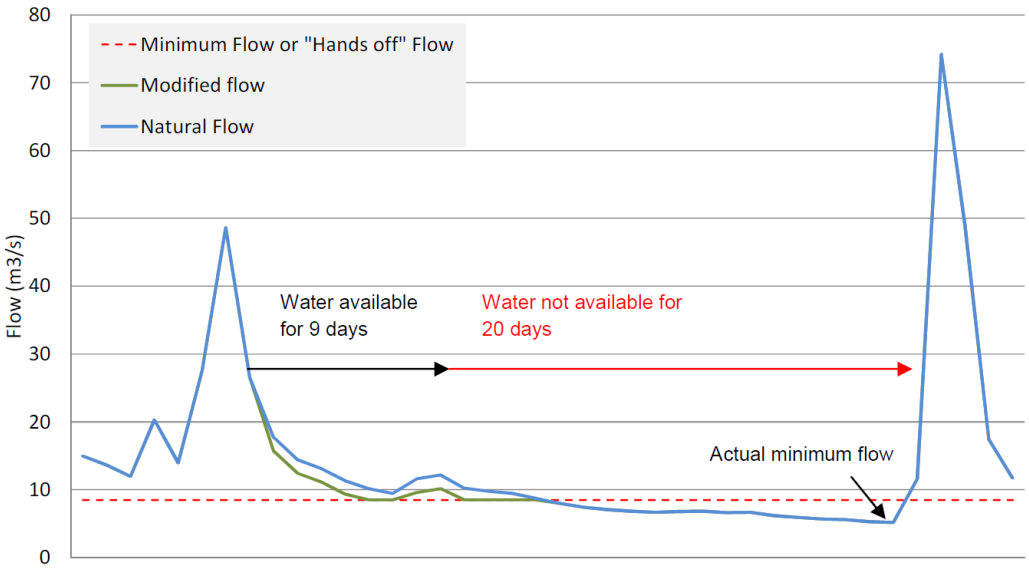



Figure 1. 3 An example of the impacts of regulations on run-of-river water availability. The red dashed line refers to a set flow (minimum flow) when the water harvesting should cease. The blue line refers to the natural flow, and the green line refers to the water flow with water allocation. The water availability is determined by the size of the water allocation and the threshold of minimum water flow. The x-axis indicates the variation in days.
Source: McLea & Thompson, 2016

## 1.2 Problem situation

### *1.2.1 The Relationships and Importance of Plant Water Status to Grape Quality*

To achieve consistent wine styles, moderate water deficit management is a common practice applied in vineyards for improving grape composition (Des Gachons et al., 2005; Pellegrino et al., 2005; Seguin, 1986; Smart, 1974; Van Leeuwen et al., 2009; Zufferey et al., 2017). The water stress induced by water deficit reduces vegetative growth and impairs photosynthesis (Hsiao, 1973; Keller, 2005), which reduces berry size, improves the partitioning of carbohydrates to clusters, alters canopy microclimate, and increases the exposure of fruit to sunlight (Keller et al., 2016; Ollat et al., 2002). This results in the synthesis of secondary metabolites and higher concentrations of quality-determining compounds, leading to better grape quality and complex aromas during bottle aging (Morrison & Noble, 1990; Ojeda et al., 2002). Positive impacts, improvement in colors, flavor, and aroma, from controlled water deficit on grape composition, due to increasing biosynthesis of phenolic compounds, anthocyanin, skin-to-pulp ratio, aroma precursors (Acevedo-Opazo et al., 2010; Des Gachons et al., 2005; Koundouras et al., 2006; Santesteban et al., 2011; Van Leeuwen et al., 2004). Water deficit, in general, is beneficial to viticulture in terms of grape profiles and wine attributes through striking the balance between carbon fixation and water loss between fruit set and veraison (Fernandes-Silva et al., 2019; Schreiner & Lee, 2014).

The management of water deficit practice in vineyards is known as deficit irrigation (DI), which includes sustained deficit irrigation (deficit through the season), regulated irrigation (irrigation is stopped or reduced based on the physiological

response of crops to water deficit), and partial root drying (application of reduced irrigation to alternate sides of the vine) (Dry et al., 2001; Fernandes-Silva et al., 2019). DI is a viable strategy to improve grape quality and has been conducted over the past two decades (Balint & Reynolds, 2013). It is carried out by applying a reduced amount of water compared to full irrigation to develop mild dehydration, with minimal effects on carbon assimilation, yield, and quality. DI strategy, generally, is at a mild level for white varieties (e.g., Sauvignon Blanc and Chardonnay) and a higher level of water stress for red varieties (e.g., Pinot Noir and Merlot) of grape, respectively (Des Gachons et al., 2005; Netzer et al., 2019). This difference originates from the different production goals, where the general tendency under increased dehydration (but not severe situations) is better for the quality of the red variety but would jeopardize the quality of the white variety (Van Leeuwen et al., 2019). The phenology of grapevine is shown in Figure 1.4. In common practice, grapevines are not irrigated until they reach moderate stress (-1200 or -1300 kilopascal (kPa)) after fruit set for the red variety (Baeza et al., 2019). For the white variety, irrigation thresholds are set to ensure no production loss, and the stress state of grapevines is brought down to near stress (-900 kPa) after veraison (Krasnow et al., 2019). Midday water stress close to -1400 kPa is avoided, as it imposes detrimental effects on photosynthesis, yield, and quality.

Either excessive water stress or over-irrigation has been reported to compromise grape quality, by inhibiting flavor development processes in the late season (Pellegrino et al., 2014; Williams et al., 2009). Questions around the exact levels of water deficit and timing of water restriction that, could have positive effects on quality-determining compounds remain unsolved (Baeza et al., 2019; Keller et al., 2016; Koundouras et al., 1999; Peterlunger et al., 2002; Scholasch & Rienth, 2019). One disadvantage of DI is that it requires plant water status to be maintained within a rather narrow range, any excess or under application would be harmful to the advantages of DI (Jones, 2004; Mirás-Avalos & Intrigliolo, 2017; Van Leeuwen & Seguin, 2006). This is a trade-off between yield losses, quality gains, and water use, which needs to be optimized according to the commercial strategies and economic return. The requirement of DI management highlights the importance of GWS monitoring, which is necessary for accurately regulating the level of water stress imposed on vines, while avoiding irreversible damage and yield loss (Acevedo-Opazo et al., 2010; Ojeda et al., 2002; Pellegrino et al., 2005).

Under the scenarios of global warming, reduced rainfall, and prolonged drought, water scarcity is foreseeable in the future. Heterogeneous vineyards limit the efficient use of uniform irrigation. As a result, there is a need for the viticultural industry to improve the water use efficiency of grapevines. In addition, competition between agricultural sectors, municipal demands, and other industries, accompanied by increasing environmental regulations, further elevates the pressure of using water resources in an efficient way. In this context, it will become progressively more difficult to reach the required balance for grape composition and expected wine styles. Therefore, it is desirable to develop and validate a tool that can assist with assessing GWS through the growing season at a large scale, determining thresholds and timing for a desired water deficit level, before triggering irrigation. To this end, precision irrigation scheduling with the ability to estimate levels of plant water stress is one of the tools that can help the viticultural industry to achieve improved DI management and optimal grape quality (Scholasch & Rienth, 2019; Stagakis et al., 2012).

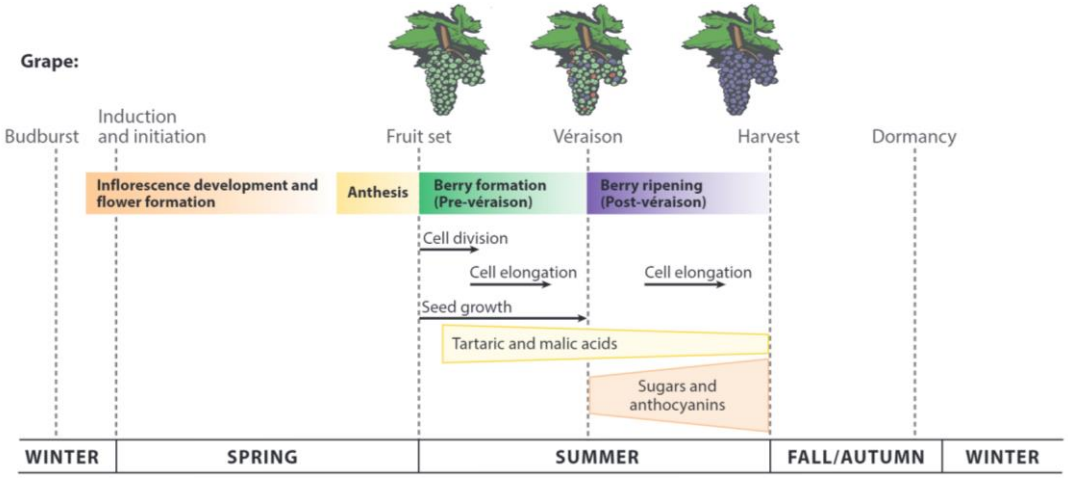



Figure 1. 4 Grapevine phenological stages including vegetative and reproductive development during the growing season.
Source: Naidu et al., 2015

### *1.2.2 The Utilization of Geospatial Techniques and Remote Sensing Information in Precision Agriculture*

To precisely monitor the response of grapevines to water stress for quality optimization in response to negative weather conditions, it is necessary to move from traditional uniform management to precision irrigation. This approach relies on the utilization of emerging technologies capable of accurately assessing the heterogeneous characteristics of water deficits in a spatial manner. Practical implementation of achieving precision irrigation includes several different stages: data collection, data analysis, and decision-making or variable rate control (Bhakta et al., 2019). This study focuses on researching data acquisition, based on the use of satellite/ airborne/ ground-based platforms with sensors and global navigation satellite systems (GNSS), and data analysis, based on the use of geographical information system (GIS) and machine learning (ML) algorithms. These technologies enable the visualizing, diagnosing, and quantifying of plant dehydration variability across the fields, which help develop optimized solutions for GWS estimation across vineyards along growing seasons.

#### 1.2.2.1 Remote Sensing and Proximal Sensing

Remote sensing (RS) allows for acquiring enormous surrogate information about an object remotely or at a certain distance (Pinter Jr et al., 2003). RS has gained popularity in precision irrigation as it enables the site-specific evaluation of plant water stress dynamics for a large coverage in a timely and non-invasive way (Matese et al., 2018; Ullah et al., 2012). In PV, RS is mainly based on reflectance estimates, the proportion of radiation striking an object's surface (irradiance) to the radiation reflected off (radiance), which is captured at different wavelengths (spectral bands) (Matese et al., 2015). The underlying principle is that the incident radiations are absorbed, transmitted, and reflected when interacting with the object's surface, and the reflectance at each spectral band is associated with various physical and chemical characteristics of the object (Wahabzada et al., 2015). Therefore, the reflectance properties are specific to the state of the object and are commonly known as spectral

signatures. As the plant water stress is the integrative response to physiological state, weather conditions, soil and terrain properties, and seasonality, it is expected that electromagnetic radiation can reflect the variability of GWS and express it through the variability of spectral signatures (Gutierrez et al., 2010; Penuelas et al., 1993).
RS typically can be classified according to the type of platform (ground, airborne, or satellite), spectral region (visible, infrared, and microwave), number and width of the spectral bands (multispectral and hyperspectral), spatial resolution (<0.1 m to 1000 m), temporal resolution (hourly, daily and weekly revisiting frequency), the source of radiation (solar-based passive or artificial light based active sensors) used to illuminate the objects, and type of signal the sensor can capture (sun-induced fluorescence, reflectance, and thermal emission). Ground-based hyperspectral reflectance and airborne/ satellite based multispectral reflectance spans across visible to infrared bands are the focus of this study.

#### 1.2.2.2 Global Navigation Satellite Systems (GNSS)

The accurate spatial location of any features or objects at any time is a fundamental part of PV. GNSS is a satellite network that allows ground receivers to calculate their precise geographic coordinates in three dimensions (including altitude, longitude, and latitude) by gauging their distance from multiple satellites (Mitran et al., 2021), providing a spatial reference to remotely sensed images collected over the fields via georeferencing process. The georeferencing process associates spatial or geographic information with forms of spatial data, so that the spatial data can be accurately positioned and aligned within a defined geographic reference system. Generally is carried out by co-register, including scaling and transformation, the aerial images with pre-surveyed coordinate points, known as ground control points (Eugenio & Marqués, 2003). GNSS also has been incorporated into a growing number of ground-based and aerial devices recently to provide geo-referenced measurements, which makes the comparison possible between different spatial data. Moreover, real-time kinematic (RTK) corrects the spatial accuracy of GNSS to around 0.02 m using a ground-base station with reference coordinates, which is desired for the mapping in vertical-oriented vineyards due to the features of discontinuous vegetation (Tisseyre et al., 2008).

#### 1.2.2.3 Geographical Information System (GIS)

GIS is a computer-based system designed to capture, store, manage, analyze, and visualize geographic data or information. It allows users to work with spatial data. GIS technology enables the integration of various data layers (e.g., maps and satellite imagery) and the ability to perform complex spatial analysis and modeling, making it a valuable tool for decision-making in fields such as natural resource management, and environmental science (Green, 2012; Supuwiningsih & Rusli, 2017).

#### 1.2.2.4 Machine learning (ML) Algorithms

Analytic approaches are needed to identify the important variables or vegetation indices related to the changes in GWS, to determine the relationships between various spatial data processed by GIS. These approaches broadly can be categorized as physics-based methods and empirical methods (Holloway & Mengersen, 2018). Physics-based methods rely on a comprehensive understanding of the underlying relationships or

physical principles governing the systems being simulated. For instance, techniques like radiative transfer models and energy balance models fall into this category. These methods conduct simulations by incorporating detailed information about the model's structure and the characteristics of its components. As a result, they bring domain-specific knowledge and physical constraints into the modeling process. Nevertheless, the availability of data on relevant components would be a concern in model development. The empirical methods establish the complex relationships between variables without explicitly formalizing the processes and factors involved in the system, so it is generally simpler. However, the results of the empirical methods depend on the quality of predictors and representativeness of ground-truth data. These make the simulated relationships sensitive to the condition of data collection, including uncertainties associated with device noise, crop physiological state, and environmental dynamics. These methods are site-dependent and unwise to extrapolate to other contexts.

A vineyard is a complex ecosystem where canopies interact with several factors (weather patterns, genotype characteristics, underground properties, and cultivation practices). Therefore, it would be difficult to carry out GWS estimation using physics-based methods as it is hard to consider all environmental factors that are associated with the changes in GWS in vineyards, and it is costly to collect the spatial dynamics of every component during the growing season. ML models, using the empirical method, have become popular in the field of agriculture (Priya & Ramesh, 2020), undertaking calculations based on the structures and patterns in the dataset. ML approaches provide a mathematically rigorous way (sampling, model training, performance evaluation, and generalization) to estimate and predict the outcomes of interest while quantifying the associated uncertainty.

Although, innovative geospatial technologies are available for incorporation into GWS monitoring systems, the application and adoption of them by the viticulture sector face many challenges. Bhakta et al. (2019) and Mitran et al. (2021), said these challenges can be classified as: organization-related, farm-related, data-related, and technology-related. This study is specially intended to use RS, GNSS, GIS, and ML algorithms to address the data-related and technology-related issues. These include:

(i) a lack of understanding of robust analysis methods for GWS-related spectral information,

(ii) a lack of data interoperability between spectral and non-spectral data obtained from various devices with different formats, and

(iii) the lack of availability of high-coverage images (larger than 10 ha), with high spatial and temporal resolution.

### 1.3 Research Aims

This research aims to address some technical and data limitations of using geospatial techniques in characterizing spatio-temporal behaviors of GWS. It is expected to provide practical approaches for improving the monitoring capacities of multispectral sensors for GWS estimation, while developing a prototype of the model

for supporting irrigation scheduling in vineyards. Additionally, this research is specific to Pinot Noir trained in the vertical-oriented system.

### 1.4 Hypotheses

Within the research scope and settings, the general hypotheses of this thesis are:

*Hypothesis 1*: The relationship between spectral data and changes in GWS can be revealed using statistical preprocessing techniques and ML algorithms.

*Hypothesis 2*: The technical shortcomings of multispectral UAVs in estimating GWS can be improved by the interoperability between spatial data collected from different sources (satellite, UAV, spectroscopy, EM38, weather station), GNSS, GIS, and ML models.

*Hypothesis 3*: The reliable prediction of GWS at the large scale can be achieved by the application of geospatial processing and modeling along with ML algorithms.

### 1.5 Objectives

To achieve the research aim stated in section 1.3, objectives are set as follows:

*Objective 1*: Determine important spectral regions relevant to GWS variation as the foundation of applying spectral information to estimating GWS.

*Objective 2*: Explore which environmental variables (temporal trends, weather conditions, and soil/terrain data) can improve the accuracy of multispectral UAV-based GWS estimation, while giving insights into the relationships and contributions that the environmental variables make toward GWS variation.

*Objective 3*: Develop a prediction model based on satellite images that can estimate and predict the variation of GWS at the large scale.

### 1.6 Thesis Outline

This thesis is divided into six chapters and contains three international peer-reviewed journal papers. Chapter 1 serves as an introduction to the research background and settings. Chapter 2 provides a review of the relevant literature. Each paper is presented as an independent chapter in Chapters 3-5. Chapter 6 describes the contribution of each paper to the overall research aim. The thesis is written in U.S. English.

**Chapter 1** – It provides the study context about the importance of GWS monitoring in the industrial, climatic, regulatory, and management context. Besides, the current state of applying multispectral RS information, GNSS, GIS, and ML models to scheduling irrigation for viticulture, while identifying the research gap - technical limitations that constrain the wide distribution of these technologies.

**Chapter 2** – It is intended to give sufficient background information on this research, starting with a general introduction to irrigation management in vineyards. It is

followed by some description of the environmental factors that drive the variation of GWS in fields. Subsequently, details are given to elaborate on ML models and statistical evaluation methods used in this research.

**Chapter 3** – This chapter investigates various modeling pipelines using hyperspectral information to provide a holistic evaluation of their performance on GWS estimation. In addition, important spectral regions are identified in accordance with these modeling pipelines (objective 1).

**Chapter 4** – This chapter aims to improve the estimating capacities of multispectral UAVs for GWS monitoring by including environmental variables from different sources, to complement the ML regression modeling. Several vegetation indices are tested against GWS variation, and hierarchical clustering is used to generate different combinations of inputs, for selecting important environmental variables. Moreover, SHapley Additive exPlanations analysis is employed to determine the contributions and relationships of the environmental variable with the changes in GWS (objective 2).

**Chapter 5** – This chapter uncovers the potential of using two-stage calibration to generate calibrated satellite images that robustly estimate ground-truth information at a larger scale. Given the calibrated satellite images, a GWS prediction tool is developed based on ML models for predicting GWS changes over the study vineyards during the critical periods when coupled with the weather forecasts (objective 3).

**Chapter 6** – It summarizes the main findings and potential implications of this research. It also provides suggestions together with some perspectives formulated for future research.

**Reference**


Acevedo-Opazo, C., Ortega-Farias, S., & Fuentes, S. (2010). Effects of grapevine (Vitis vinifera L.) water status on water consumption, vegetative growth and grape quality: An irrigation scheduling application to achieve regulated deficit irrigation. *Agricultural Water Management, 97*(7), 956-964.

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*(4), 308.

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*.

Baeza, P., Junquera, P., Peiro, E., Lissarrague, J. R., Uriarte, D., & Vilanova, M. (2019). Effects of vine water status on yield components, vegetative response and must and wine composition. *Advances in Grape and Wine Biotechnology*.

Balafoutis, A. T., Koundouras, S., Anastasiou, E., Fountas, S., & Arvanitis, K. (2017). Life cycle assessment of two vineyards after the application of precision viticulture techniques: A case study. *Sustainability, 9*(11), 1997.

Balint, G., & Reynolds, A. G. (2013). Effect of different irrigation strategies on vine physiology, yield, grape composition and sensory profile of Sauvignon blanc (Vitis vinifera L.) in a cool climate area. *J. Int. Sci. Vigne Vin, 47*(3), 159-181.

Bhakta, I., Phadikar, S., & Majumder, K. (2019). State-of-the-art technologies in precision agriculture: a systematic review. *Journal of the Science of Food and Agriculture, 99*(11), 4878-4888.

Bramley, R. (2022). Precision Viticulture: Managing vineyard variability for improved quality outcomes. In *Managing wine quality* (pp. 541-586). Elsevier.

Bramley, R., & Lamb, D. (2003). Making sense of vineyard variability in Australia. Proc. Internat. Symp. on Precision Viticulture, Ninth Latin American Congr. on Viticulture and Oenology,

Bramley, R., Proffitt, A., Hinze, C., Pearse, B., & Hamilton, R. (2005). Generating benefits from precision viticulture through selective harvesting. *Precision Agriculture, 5*, 891-898.

Des Gachons, C. P., Leeuwen, C. V., Tominaga, T., Soyer, J. P., Gaudillère, J. P., & Dubourdieu, D. (2005). Influence of water and nitrogen deficit on fruit ripening and aroma potential of Vitis vinifera L cv Sauvignon blanc in field conditions. *Journal of the Science of Food and Agriculture, 85*(1), 73-85.

Dry, P. R., Loveys, B. R., McCarthy, M. G., & Stoll, M. (2001). Strategic irrigation management in Australian vineyards. *OENO One, 35*(3), 129. https://doi.org/10.20870/oeno-one.2001.35.3.1699

Eugenio, F., & Marqués, F. (2003). Automatic satellite image georeferencing using a contour-matching approach. *IEEE Transactions on Geoscience and Remote Sensing, 41*(12), 2869-2880.

Fernandes-Silva, A., Oliveira, M., Paço, T. A., & Ferreira, I. (2019). Deficit irrigation in Mediterranean fruit trees and grapevines: Water stress indicators and crop responses. *Irrigation in Agroecosystems. London, United Kingdom*.

Fraga, H., & Santos, J. A. (2017). Daily prediction of seasonal grapevine production in the Douro wine region based on favourable meteorological conditions. *Australian Journal of Grape and Wine Research, 23*(2), 296-304.

Green, D. R. (2012). Geospatial tools and techniques for vineyard management in the twenty-first century. In *The geography of wine* (pp. 227-245). Springer.

Gutierrez, M., Reynolds, M. P., & Klatt, A. R. (2010, Jul). Association of water spectral indices with plant and soil water relations in contrasting wheat genotypes. *J Exp Bot, 61*(12), 3291-3303. https://doi.org/10.1093/jxb/erq156

Holloway, J., & Mengersen, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: a review. *Remote Sensing, 10*(9), 1365.

Hsiao, T. C. (1973). Plant responses to water stress. *Annual review of plant physiology, 24*(1), 519-570.

IPCC. (2014). *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change*.

Jackson, D., & Lombard, P. (1993). Environmental and management practices affecting grape composition and wine quality-a review. *American Journal of Enology and Viticulture, 44*(4), 409-430.

Jones, G. (2006). Climate change and wine: observations, impacts and future implications. *Wine Industry Journal, 21*(4), 21-26.

Jones, G. (2021). Wine Production and Climate Change. In *World Scientific Encyclopedia of Climate Change*. World Scientific.

Jones, G. V. (2018). The climate component of terroir. *Elements: An International Magazine of Mineralogy, Geochemistry, and Petrology, 14*(3), 167-172.

Jones, H. G. (2004, Nov). Irrigation scheduling: advantages and pitfalls of plant-based methods. *J Exp Bot, 55*(407), 2427-2436. https://doi.org/10.1093/jxb/erh213

Keller, M. (2005). Deficit irrigation and vine mineral nutrition. *American Journal of Enology and Viticulture, 56*(3), 267-283.

Keller, M., Romero, P., Gohil, H., Smithyman, R. P., Riley, W. R., Casassa, L. F., & Harbertson, J. F. (2016). Deficit irrigation alters grapevine growth, physiology, and fruit microclimate. *American Journal of Enology and Viticulture, 67*(4), 426-435.

King, P. D., Smart, R., & McClellan, D. J. (2014). Within-vineyard variability in vine vegetative growth, yield, and fruit and wine composition of C abernet S auvignon in H awke's B ay, N ew Z ealand. *Australian Journal of Grape and Wine Research, 20*(2), 234-246.

Koundouras, S., Marinos, V., Gkoulioti, A., Kotseridis, Y., & van Leeuwen, C. (2006). Influence of vineyard location and vine water status on fruit maturation of nonirrigated cv. Agiorgitiko (Vitis vinifera L.). Effects on wine phenolic and aroma components. *Journal of Agricultural and Food Chemistry, 54*(14), 5077-5086.

Koundouras, S., Van Leeuwen, C., Seguin, G., & Glories, Y. (1999). Influence of water status on vine vegetative growth, berry ripening and wine characteristics in mediterranean zone (example of Nemea, Greece, variety Saint-George, 1997). *OENO One, 33*(4), 149-160.

Krasnow, M., Haywood, A., & McMillan, D. (2019). Irrigation optimization.

Langemeier, M., & Boehlje, M. (2021). What Will Be the Capabilities and Skills Needed to Manage the Farm of the Future? *farmdoc daily, 11*(61).

Levin, A. D., Williams, L. E., & Matthews, M. A. (2020). A continuum of stomatal responses to water deficits among 17 wine grape cultivars (Vitis vinifera). *Functional Plant Biology, 47*(1), 11-25.

Matese, A., Baraldi, R., Berton, A., Cesaraccio, C., Di Gennaro, S. F., Duce, P., Facini, O., Mameli, M. G., Piga, A., & Zaldei, A. (2018). Estimation of water stress in grapevines using proximal and remote sensing methods. *Remote Sensing, 10*(1), 114.

Matese, A., Toscano, P., Di Gennaro, S., Genesio, L., Vaccari, F., 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*(3), 2971-2990. https://doi.org/10.3390/rs70302971

McLea, M., & Thompson, M. (2016). *Concepts for managing water allocation*.

MfE. (2018). *Climate Change Projections for New Zealand: Atmosphere Projections Based on Simulations from the IPCC Fifth Assessment*.

MfE. (2021). *Measuring and reporting water takes*.

Mirás-Avalos, J. M., & Intrigliolo, D. S. (2017). Grape composition under abiotic constrains: water stress and salinity. *Frontiers in Plant Science, 8*, 851.

Mitran, T., Meena, R. S., & Chakraborty, A. (2021). Geospatial Technologies for Crops and Soils: An Overview. *Geospatial Technologies for Crops and Soils*, 1-48.

Morrison, J. C., & Noble, A. C. (1990). The effects of leaf and cluster shading on the composition of Cabernet Sauvignon grapes and on fruit and wine sensory properties. *American Journal of Enology and Viticulture, 41*(3), 193-200.

Naidu, R. A., Maree, H. J., & Burger, J. T. (2015). Grapevine leafroll disease and associated viruses: a unique pathosystem. *Annual Review of Phytopathology, 53*(1), 613-634.

Neethling, E., Petitjean, T., Quénol, H., & Barbeau, G. (2017). Assessing local climate vulnerability and winegrowers' adaptive processes in the context of climate change. *Mitigation and Adaptation Strategies for Global Change, 22*(5), 777-803.

Netzer, Y., Munitz, S., Shtein, I., & Schwartz, A. (2019). Structural memory in grapevines: Early season water availability affects late season drought stress severity. *European Journal of Agronomy, 105*, 96-103. https://doi.org/10.1016/j.eja.2019.02.008

OIV. (2022). STATE OF THE WORLD VINE AND WINE SECTOR 2021. https://www.oiv.int/sites/default/files/documents/eng-state-of-the-world-vine-and-wine-sector-april-2022-v6_0.pdf

Ojeda, H., Andary, C., Kraeva, E., Carbonneau, A., & Deloire, A. (2002). Influence of pre- and postveraison water deficit on synthesis and concentration of skin phenolic compounds during berry growth of Vitis vinifera cv. Shiraz. *American Journal of Enology and Viticulture, 53*(4), 261-267.

Oliver, M. A., Bishop, T. F., & Marchant, B. P. (2013). *Precision agriculture for sustainability and environmental protection*. Routledge Abingdon.

Ollat, N., Carde, J.-P., Gaudillère, J.-P., Barrieu, F., Diakou-Verdin, P., & Moing, A. (2002). Grape berry development: a review. *OENO One, 36*(3), 109-131.

Onibonoje, M. O., & Nwulu, N. (2021). Synergistic technologies for precision

agriculture. In *Artificial Intelligence and IoT-Based Technologies for Sustainable Farming and Smart Agriculture* (pp. 123-139). IGI global.

Pellegrino, A., Clingeleffer, P., Cooley, N., & Walker, R. (2014). Management practices impact vine carbohydrate status to a greater extent than vine productivity. *Frontiers in Plant Science, 5*, 283.

Pellegrino, A., Lebon, E., Simonneau, T., & Wery, J. (2005). Towards a simple indicator of water stress in grapevine (Vitis vinifera L.) based on the differential sensitivities of vegetative growth components. *Australian Journal of Grape and Wine Research, 11*(3), 306-315.

Penuelas, J., Gamon, J. A., Griffin, K. L., & Field, C. B. J. R. S. o. E. (1993). Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. *46*(2), 110-118.

Peterlunger, E., Sivilotti, P., Bonetto, C., & Paladin, M. (2002). Water stress induces changes in polyphenol concentration in Merlot grapes and wines. *Rivista de Viticoltura e di Enologia, 1*, 51-66.

Pinter Jr, P. J., Hatfield, J. L., Schepers, J. S., Barnes, E. M., Moran, M. S., Daughtry, C. S., & Upchurch, D. R. (2003). Remote sensing for crop management. *Photogrammetric Engineering and Remote Sensing, 69*(6), 647-664.

Plant & Food Research. (2021). NEW ZEALAND HORTICULTURAL EXPORTS. https://www.freshfacts.co.nz/files/freshfacts-2021.pdf

Priya, R., & Ramesh, D. (2020). ML based sustainable precision agriculture: A future generation perspective. *Sustainable Computing: Informatics and Systems, 28*, 100439.

Proffitt, A. P. B., Bramley, R., Lamb, D., & Winter, E. (2006). *Precision viticulture: a new era in vineyard management and wine production*. Ashford, Winetitles.

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. https://doi.org/10.1016/j.compag.2018.02.013

Santesteban, L., Miranda, C., & Royo, J. (2011). Regulated deficit irrigation effects on growth, yield, grape quality and individual anthocyanin composition in Vitis vinifera L. cv.'Tempranillo'. *Agricultural Water Management, 98*(7), 1171-1179.

Santos, J. A., Fraga, H., Malheiro, A. C., Moutinho-Pereira, J., Dinis, L.-T., Correia, C., Moriondo, M., Leolini, L., Dibari, C., & Costafreda-Aumedes, S. (2020). A review of the potential climate change impacts and adaptation options for European viticulture. *Applied Sciences, 10*(9), 3092.

Sassu, A., Ghiani, L., Salvati, L., Mercenaro, L., Deidda, A., & Gambella, F. (2021). Integrating UAVs and Canopy Height Models in Vineyard Management: A Time-Space Approach. *Remote Sensing, 14*(1), 130.

Scholasch, T., & Rienth, M. (2019). Review of water deficit mediated changes in vine and berry physiology; Consequences for the optimization of irrigation strategies. *OENO One, 53*(3).

Schreiner, R. P., & Lee, J. (2014). Effects of post-véraison water deficit on 'Pinot noir'yield and nutrient status in leaves, clusters, and musts. *HortScience, 49*(10), 1335-1340.

Seguin, G. (1986). 'Terroirs' and pedology of wine growing. *Experientia, 42*(8), 861-873.

Singh, A. P., Yerudkar, A., Mariani, V., Iannelli, L., & Glielmo, L. (2022). A Bibliometric Review of the Use of Unmanned Aerial Vehicles in Precision Agriculture and Precision Viticulture for Sensing Applications. *Remote Sensing, 14*(7), 1604.

Smart, R. (1974). Aspects of water relations of the grapevine (Vitis vinifera). *American Journal of Enology and Viticulture, 25*(2), 84-91.

Solomon, S., Qin, D., Manning, M., Averyt, K., & Marquis, M. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge university press.

Stagakis, S., González-Dugo, V., Cid, P., Guillén-Climent, M. L., & Zarco-Tejada, P. J. (2012). Monitoring water stress and fruit quality in an orange orchard under regulated deficit irrigation using narrow-band structural and physiological remote sensing indices. *ISPRS Journal of Photogrammetry and Remote Sensing, 71*, 47-61. https://doi.org/10.1016/j.isprsjprs.2012.05.003

Stocker, T. (2014). *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge university press.

Supuwiningsih, N. N., & Rusli, M. (2017). Prediction of decreasing agricultural land based on geographic information system case study: Denpasar city. *International Journal of Computer Application, 162*(9), 30-34.

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*(1), 27-39.

Tisseyre, B., & Taylor, J. (2005). An overview of methodologies and technologies for implementing precision agriculture in viticulture. XII Congresso Brasileiro de Viticultura e Enologia Anais,

Ullah, S., Schlerf, M., Skidmore, A. K., & Hecker, C. (2012). Identifying plant species using mid-wave infrared (2.5–6μm) and thermal infrared (8–14μm) emissivity spectra. *Remote Sensing of Environment, 118*, 95-102. https://doi.org/10.1016/j.rse.2011.11.008

Van Leeuwen, C., & Darriet, P. (2016). The impact of climate change on viticulture and wine quality. *Journal of Wine Economics, 11*(1), 150-167.

Van Leeuwen, C., Destrac-Irvine, A., Dubernet, M., Duchêne, E., Gowdy, M., Marguerit, E., Pieri, P., Parker, A., De Resseguier, L., & Ollat, N. (2019). An update on the impact of climate change in viticulture and potential adaptations. *Agronomy, 9*(9), 514.

Van Leeuwen, C., Friant, P., Chone, X., Tregoat, O., Koundouras, S., & Dubourdieu, D. (2004). Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and Viticulture, 55*(3), 207-217.

Van Leeuwen, C., & Seguin, G. (2006). The concept of terroir in viticulture. *Journal of Wine Research, 17*(1), 1-10.

Van Leeuwen, C., Trégoat, O., Choné, X., Bois, B., Pernet, D., & Gaudillère, J.-P. (2009). Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? *OENO One, 43*(3), 121-134.

Wahabzada, M., Mahlein, A. K., Bauckhage, C., Steiner, U., Oerke, E. C., & Kersting, K. (2015). Metro maps of plant disease dynamics--automated mining of differences using hyperspectral images. *PLoS One, 10*(1), e0116902. https://doi.org/10.1371/journal.pone.0116902

Williams, L. E., Grimes, D. W., & Phene, C. J. (2009). The effects of applied water at various fractions of measured evapotranspiration on reproductive growth and water productivity of Thompson Seedless grapevines. *Irrigation Science, 28*(3), 233-243. https://doi.org/10.1007/s00271-009-0173-0

Winegrowers, N. Z. (2022). Vineyard Report https://www.nzwine.com/media/22001/1-vineyard-report-2022.pdf

Zufferey, V., Spring, J.-L., Verdenal, T., Dienes, A., Belcher, S., Lorenzini, F., Koestel, C., Rösti, J., Gindro, K., & Spangenberg, J. (2017). The influence of water stress on plant hydraulics, gas exchange, berry composition and quality of Pinot Noir wines in Switzerland. *OENO One, 51*(1).

# Chapter 2
# Literature Review

## 2.1 References of Plant Water Status

### *2.1.1 Mechanisms behind Plant Water Stress*

Measuring GWS for carrying out precision irrigation requires a better understanding of the morphological and physiological mechanisms of plant responses to water stress. Plants absorb soil water around root zones to meet their evapotranspiration (ET) needs. Under the conditions of limiting soil moisture or high atmospheric demand, leaf curling due to loss of cell turgidity may take place to reduce exposed leaf area and water loss through stomata (Kadioglu et al., 2012). In addition, chemical and hydraulic signals are induced and transmitted to the leaves through xylem pathways, leading to stomatal closure and reduced transpiration (Limpus, 2009). Consequently, the photosynthesis rate is decreased since CO2 uptake and fixation are inhibited (Zhou et al., 2007). The reduced photosynthesis induces oxidative stress and then results in decreasing chlorophyll content, which is indicated by the occurrence of yellow leaves (Gholamin & Khayatnezhad, 2011). A persistent water deficit will further damage photosynthetic machinery through loss of chlorophyll and electrolyte leakage, which slows down cell division with wilting leaves, reduced leaf area, and vegetative growth (Baher et al., 2002; Massonnet et al., 2007). In the reproductive growth stage, flower abortion and cluster abscission may be observed under dehydration, possibly associated with changes in hormones, which causes reduced biomass, yield, and quality of crops (Matthews & Anderson, 1989; Rossini et al., 2013).

The duration, intensity, and timing of water stress determine the responses and impacts on plant metabolism (Sun et al., 2017). For example, under mild water deficit, plants maintain a certain degree of leaf water content without or with only a few effects on photosynthetic capacity through regulation of water loss and uptake. On the contrary, severe water stress causes a great extent of physiological and biochemical changes, which induces photosynthetic impediment and thus stunts the growth of crops. Water status monitoring hence plays a key role in farm management for crop performance. Although visual field observation is one of the simple ways to assess plant water status, irreversible damage can take place before the appearance of visible symptoms of water stress (Jones & Schofield, 2008). Therefore, a pre-symptomatic approach for early detection of plant water status can be employed to avoid severe crop damage while increasing water use efficiency.

### *2.1.2 Conventional Methods for Determining Plant Water Stress*

Precise estimates of plant water status are helpful for decision-making (when and where to apply) in irrigation scheduling (Peñuelas et al., 1997). Traditional methods for describing water stress rely on in-situ continuous measurement of soil water content using soil moisture sensors (Thompson et al., 2007) or the estimated water loss (ET) from the plant-soil systems during a given period provided by local agricultural agencies based on the observed meteorological parameters (Smith & Munoz, 2002). Other methods involving direct or indirect plant-based measurements, such as sap flow, trunk diameter, stomatal conductance, and plant water potential, also have been

used in assessing plant water deficit (Ihuoma & Madramootoo, 2017). Plant-based methods use plants as biosensors, combining the effects of soil moisture content, atmospheric demands, internal hydraulic resistance, uptake capacity of the plant-root interface, and plant physiological responses to water availability, to present a holistic assessment of water stress level (Jones & Vaughan, 2010).

One important drawback of using those conventional methods for detecting water stress is the flaws in the assumption of uniformity. For soil moisture sensors, it is challenging to characterize the spatial distribution, from horizontal and vertical perspectives, of water retention characteristics across the entire field based only on a few measurement points. Heterogeneous plots in terms of soil properties are often encountered in viticulture. Besides, it is common to deliver water in vineyards through drip irrigation rather than broadcast or sprinkler irrigation. The nature of drip irrigation imposes extra interference on the soil water content in the crop rows compared to the inter-rows, which makes extrapolation of vine water status from the soil moisture sensed in the inter-rows rather challenging. Ideally, complete soil mapping needs to be implemented prior to the installation of soil sensors, followed by adjusting the number and location of sensors according to soil heterogeneity. However, it is usually a financial burden to acquire representative measurements in vineyards with diverse soil characteristics due to the requirement of a high number of sensors to account for the heterogeneity (Jones, 2012). Additionally, most soil moisture probes cannot capture the water availability to the crops which is perceived by the deep root systems, further limiting the possibilities of soil-based methods (Rienth & Scholasch, 2019). For example, grapevine is capable of accessing water up to 9 m while customer-grade soil moisture probes generally extend to 1.5 m of depth (Gautam & Pagay, 2020). Similar conditions are applied to ET models, as they assume uniform vegetation cover, freely transpiring in a ratio to the reference crop. Moreover, soil-based, and atmospheric approaches give an assessment of the imposed water deficit, instead of representing the level of water stress actually experienced by the plants (Sinclair & Ludlow, 1985). It is known that plant water status is not simply connected to soil water measurements in the root zone (Lavoie-Lamoureux et al., 2017). The same degree of plant water status can happen at several levels of soil water availabilities since there are uncertainties about where the active root zones, responsible for major water uptake, are located (Davenport et al., 2008; Jones, 2004). This makes it necessary to consider other environmental factors when applying soil moisture sensors to schedule irrigation. Although the measurements of sap flow and plant water potential provide direct and robust information about plant water status, their requirement for plant contact or destructive sampling makes it featured as labor-intensive, time-consuming in acquiring a reasonable amount of data, unsuitable for automation and continuous monitoring, and limiting applicability (Meron et al., 1987; Santos & Kaye, 2009). Like soil-based and atmospheric approaches, these plant-based methods can only provide point information as a reference of average plant water status for guiding the irrigation scheduling across entire fields. As a result, it is foreseeable that over- or under-irrigation can occur when using conventional methods for monitoring plant water deficit, especially for the severely non-homogeneous blocks. This study used plant water potential, stem water potential specifically, as the ground-truth information of GWS to calibrate remotely or proximally sensed data.

### *2.1.3 Plant Water Potential*

In modern vineyards, water potential (Ψ) in vines derived from a pressure chamber is frequently used indicator to evaluate the GWS due to mechanical simplicity, low maintenance, and robustness (Scholander et al., 1965; Van Leeuwen et al., 2009). Although it is argued that the measured Ψ has already been subject to some physiological control (Jones, 2004), its integrative sensitivity to climate and soil conditions still makes it a good proxy to crop water status. As Ψ refers to the suction or the negative pressure, it is lower in plants compared to in soil to enable the absorption of water from the ground to crops. The decreasing gradient of Ψ along different parts of plants preserves a constant water flow from roots to leaves, and later water vapor transpires through the stomata. If soil water potential is reduced, the respective decrease of Ψ takes place to guarantee water supply for photosynthesis, vegetative, and reproductive growth (Scholander et al., 1965).

The method of Ψ measurement starts with placing a leaf or a leaf with the petiole (sample) immediately inside a sealed pressure bomb after sampling. The cut end on the petiole segment is required to be smooth and clean for an easier assessment of the appearance of drops later (Fernández, 2017). Pressurized air is then slowly released into the chamber to gradually increase the inside pressure. Greater rates of pressure increase would cause errors in measurement due to the rising temperature and the inertia resulting from plant hydraulic resistance. When the sap drops in the xylem are forced out by the increasing pressure and appear at the cut end, the pressure applied at this point is equal and opposite to the Ψ of the sample. According to the measurement time and sampled part, the measured Ψ can be classified as midday leaf water potential, predawn leaf water potential, and stem water potential.

#### 2.1.3.1 Midday Leaf Water Potential ($\Psi_{leaf}$)

$\Psi_{leaf}$ is the measurement of Ψ in leaf and petiole, which is usually taken at midday. The term 'midday' here means the moment of the day when minimum values of $\Psi_{leaf}$ are recorded, which is usually 2-4 hours after solar noon (Fernández, 2014). To acquire a proper measurement, well-exposed and median-size leaves are sampled from the outer layers of the canopies. In plants with tall canopies, representative leaves are usually taken at about 1.4-1.7 m aboveground (Fernández, 2017). $\Psi_{leaf}$ generally is used more often by growers because this measurement can be obtained in a faster manner (compared to stem water potential) at a convenient time of measuring (compared to predawn leaf water potential). The drawback of this tool is that it is highly influenced by the microclimate encompassing each particular leaf (Jones, 2004). It can serve as a sensitive indicator for irrigation control only under very steady climatic conditions (Williams & Araujo, 2002).

#### 2.1.3.2 Predawn Leaf Water Potential ($\Psi_{pd}$)

This measurement is usually taken place just before sunrise on adult leaves. It is assumed that the Ψ attains equilibrium overnight between crop and soil, which gives the maximum value of $\Psi_{pd}$ at predawn (Améglio et al., 1999). It is thought to be a good reflection of the soil moisture and the rehydration level that the plants have achieved overnight, which turns out to be an index of water status for crops. This has the advantage as they are not influenced by inherent stomatal behaviors during the day. Nevertheless, inaccuracy can happen to $\Psi_{pd}$ if the stomata open during the nighttime

with constant water loss, thus leading to an imbalance of Ψ between soil and crops. This phenomenon has been reported for grapevines in Mediterranean conditions (Fuentes et al., 2014). $\Psi_{pd}$ measurement is hardly applied at the commercial scale because of the need of measuring before daybreak (Rienth & Scholasch, 2019).

#### 2.1.3.3 Stem Water Potential ($\Psi_{stem}$)

$\Psi_{stem}$ is the combined result of plant transpiration and root/soil hydraulic conductivity. It has been expressed as a comprehensive and more sensitive indicator for early water deficit in vines during the day (Choné, 2001). Similar to $\Psi_{leaf}$, $\Psi_{stem}$ is usually measured in the midday when the minimum value is recorded. Since it is stable and sensitive, a few bagged leaves are enough to represent the water status of a vine (Rienth & Scholasch, 2019). The additional operation of this measurement is enclosing a sample in a plastic bag or with aluminium foil for 45-120 minutes when the samples are still attached to the canopies. In this way, transpiration is stopped with attaining equilibrium of water potential between the leaf and stem in the vicinity (Greenspan et al., 1996). Compared to $\Psi_{leaf}$, $\Psi_{stem}$ has been widely considered a better indicator of plant water status during the day, since $\Psi_{stem}$ is less influenced by the changing microclimate and thus the reading is less variable (Naor & Cohen, 2003). Under very stable environmental conditions, $\Psi_{pd}$, $\Psi_{leaf}$, and $\Psi_{stem}$ can be equally regarded as feasible methods for examining GWS (Williams & Araujo, 2002).

## 2.2 Remote and Proximal Sensing-related Terms

### *2.2.1 Platforms*

Platforms used to obtain remotely or proximally sensed information can be divided into three types: ground-based solutions, aircraft/ unmanned aerial vehicles (UAV), and satellites. Sensors can be mounted on these platforms to collect reflected radiation. Ground-based sensing uses towers, tripods, and poles as platforms to acquire ground observations or employs a moving vehicle to carry out continuous measurements. The sensor under this category is at a short distance from the objects when acquiring the reflected radiation (Ammoniaci et al., 2021). This study used a handheld, contact-based, and non-imaging (point) spectroradiometer with an artificial light source to acquire the electromagnetic radiation reflected by the sampled objects. It allows rapid collection of indirect information about the vegetation state in a non-destructive way. The active sensors, emitting radiation, enable fewer impacts from external lighting conditions, whilst contact measurement with a solid measuring angle increases the signal-to-noise ratio. However, a high amount of survey time is needed because it only provides point information on each sample object. Post-processing in generating maps by interpolating the point values is required as the collected data are tabular (reflectance value at every wavelength). To address the limitations of ground-based sensors, RS platforms are introduced as follows.

Airborne sensing acquires RS information using downward or sideward sensors. The most commonly used airborne platform in viticulture is UAVs which can be equipped with different types of sensors to assess plant status with a broad range of spectral information (Sozzi et al., 2020). UAVs are able to fly autonomously after setting up the flying route thanks to flight software, apps, embedded control units, and microprocessors. UAVs are useful in terms of flexibility, which reduces the planning time and allowing to take off whenever the weather is favored, even under cloudy

conditions. Compared to ground-based measurement, UAVs provide larger coverage of the fields in a shorter timeframe, so it makes problem detection easier and more efficient. Nevertheless, the operational endurance and monitoring area of UAVs are limited by their payload weight, and specialized post-processing procedures for UAV-based images are difficult for growers to conduct independently.

The need for monitoring and observing at a larger spatial scale has prompted the launch of several earth observation satellites (Belward & Skøien, 2015). The continuous launch of the satellite, with new sensors and improved capacities, from the same family leads to the formation of satellite constellations (e.g. Landsat, Sentinel, and WorldView families). To repeatably make use of solar radiation, they are usually sun-synchronous and their orbits are nearly passing over the poles. The technical specifications of the onboard sensors of these satellite systems are fixed after launching. Due to the larger scale, it is easy for the entire field to be sensed simultaneously by a single scene with radiometric consistency. Therefore, there is no requirement for mosaicking during the post-processing (Matese et al., 2015). Although satellite-based sensing is influenced by weather conditions, it is considered a fundamental tool for long-term crop monitoring at an extensive scale (Pastonchi et al., 2020).

### *2.2.2 Spectral Resolution*

The reflected electromagnetic energy from the sampled object reaching the sensor, through the slit and hitting the photosensitive area of the detector, is recorded at different wavelengths of the electromagnetic spectrum (Huang et al., 2001). The ability of how fine a sensor can identify the wavelength intervals (also called band or channel) of the spectrum is called spectral resolution. The spectral systems are characterized by the number, width, and position of the bands and provide a unique combination of spectral information. Based on the number of bands across the spectrum, the reflectance sensors can be classified as panchromatic (1 band), RGB (3 channels), multispectral (4-20 channels), and hyperspectral sensors (more than 20 channels) (Kozma-Bognár & Berke, 2010). Most common airborne multispectral sensors focus on blue, green, red, red edge, and near-infrared (NIR) with less spectral resolution (at least 40 nm wide) (Matese et al., 2015). Hyperspectral sensors collect reflectance data in hundreds of contiguous narrow bands (typically 10 nm) over the spectral range, which has the possibility to provide further insights into the relationship between the spectral extraction and features of interest (Pinter Jr et al., 2003). However, it was noted that enormous data sets the “curse of dimensionality” would cause a reduction in accuracy (Loggenberg et al., 2018).

### *2.2.3 Spatial Resolution*

Large field monitoring for grapevines requires sensor systems to have a high spatial resolution. The ability of a sensor to resolve the smallest spatial area and record the spectral attributes in each pixel of the image is called spatial resolution (Liang & Wang, 2019). A higher spatial resolution indicates a smaller ground sampling distance. The use of satellites in crop monitoring has great potential, but most of their fixed spatial resolutions are not sufficient for PV due to the discontinuous vegetation cover and the width of the vine rows. Satellite images with a resolution larger than 0.5 m

are considered difficult to remove inter-row pixels (Ammoniaci et al., 2021), leading to pixels with mixed spectral signals composed of canopies, ground vegetation, and/or bare soil (Khaliq et al., 2019). This decreases the effectiveness of using satellites to monitor vineyards of small sizes or irregular shapes. UAVs, on the contrary, are highly flexible in spatial resolution (down to cm) by changing the flying height. This characteristic makes UAVs ideal for providing canopy level observation and mapping intra-field variability, especially in regions with diverse parcels due to high heterogeneity, but only for vineyards of small-to-medium size (1-10 ha) due to economic feasibility (Matese & Di Gennaro, 2015).

### *2.2.4 Temporal Resolution*

Higher revisiting frequency, also known as temporal resolution, is favored because it enables satellite systems to monitor field variability with consistent spectral configurations while having temporal continuity throughout the growing seasons and across multiple years. Although most of the satellites offer a fixed temporal resolution of 1-2 weeks, a novel technological approach has been developed to reduce the time interval between satellite passes over the same location on Earth. It uses microsatellites or nanosatellites, deployed on the same orbit in a large number (20-100), to substantially enhance the number and frequency of observations collected by a satellite constellation (Toth & Jóźków, 2016). One of the first examples comes from Planet Labs which launched several “dove” satellites to test the performance of this system (Planet Team, 2017).

### *2.2.5 Spectrum*

The spectral reflectance of leaves provides information on their water content in various wavelength regions. To assess GWS using spectral information, it is necessary to identify the important fragment of the spectrum relevant to changes in plant water status. In the visible band (VIS: 400-700 nm), the pigments such as chlorophyll, carotenoids, and anthocyanin located in mesophyll are the major sources of the response. This band has the characteristics of low reflectance and high absorbance, especially in red and blue bands, while peaking in the green band (Shimada et al., 2012). It was found under water deficiency the reflectance of blue and red bands is increased while decreasing in the green band due to reduced chlorophyll concentration. The fluctuation of chlorophyll is expressed by alterations in the yellow band as well. This band is sensitive to nitrogen levels as up to 75% of the leaf nitrogen is in the chloroplasts (Evans & Seemann, 1989). Changes in chlorophyll content due to water deficit can be demonstrated by the placement and shape of the red edge band (Blackburn, 2007), and the corresponding movement of the red edge to the shorter wavelengths is known as blue shift (Campbell et al., 2007). Carotenoids are a group of pigments consisting of alpha-carotene, beta-carotene, and xanthophyll pigments. They are observed with a high concentration in plants that are stressed, senescent, or dead (Ougham et al., 2005). Anthocyanin is a reddish pigment. As it is very susceptible to oxidative damage, the concentration declines as stress ascends. It has been reported that spectral changes around 550 nm were observed in grapevines due to the changes in anthocyanin concentration (Steele et al., 2009). However, some research found that the spectral information between wavelengths of 500 and 800 nm is not relevant for

field estimation of vine water status, and they suggested that the spectra greater than 800 nm could be better for surveys instead (Rapaport et al., 2015).

High reflectance usually occurs in NIR (700-1300 nm) for green and healthy plants. Rallo et al. (2014) stated the spectra of NIR to shortwave infrared (SWIR: 1300-2700 nm) bands provided a better prediction of leaf water potential than the VIS band. NIR has two weak water absorption bands at around 970 and 1200 nm, which are engaged with the O-H bonds of water molecules (Woolley, 1971). Besides, the response in the NIR is partially due to the internal structure of the leaf, through the discontinuity in the refractive parameters between the airspace and spongy mesophyll (Gausman, 1974). Therefore, the increase of NIR reflectance under dehydration can be related to scattering by rising air content and the effects on the leaf mesophyll structure (Satterwhite & Henley, 1990). Meanwhile, water deficit results in the decomposition of celluloses and proteins in leaves and the accretion of essential amino acids and free sugars. These metabolic changes present in the NIR band, such as N-H stretches of proteins reflect around 1100 and 1400 nm, while C-H stretches of carbohydrates and proteins reflect around 1200 and 1300 nm (Yin et al., 2017). The response in the SWIR is determined partially by the nitrogen and various forms of carbon in leaves but is mostly influenced by the water content due to the existence of major water absorption bands at around 1450, 1940, and 2500 nm (Datt, 1999). Beyond SWIR, the middle infrared band (MIR) is considered insufficient in estimating leaf water status because a low signal due to strong water absorbance was observed, with insignificant reflectance changes within biologically meaningful ranges (Hunt Jr & Rock, 1989). Thus, MIR has not been extensively employed in GWS evaluation.

### *2.2.6 Orthorectification*

Orthorectification is a geometric correction process applied to remote sensing or aerial imagery to remove distortions caused by terrain relief, camera tilt, and other factors. The goal of orthorectification is to produce images with uniform scale and minimal geometric distortions so that they can be accurately used for mapping, measurement, and analysis. This process involves the precise adjustment of pixel positions in the image to align them with their corresponding geographic locations on the Earth's surface, typically using elevation data and complex mathematical transformations (Baboo & Devi, 2011). For example, topographically diverse landscapes inherit some distortions in the images. The tilt of the aerial or space-borne sensors themselves would influence the distance between objects in the images. The geometry distortions happen when the sensors are not pointing directly toward their Nadir location. Conducting this process requires an accurate terrain model and sensor model which have detailed information about the orientation and sensor location for each image.

### *2.2.7 Radiometric Correction*

Radiometric correction is a fundamental process in remote sensing and image analysis that involves adjustments to the pixel values of satellite or aerial imagery to remove variations caused by factors such as sensor characteristics, illumination conditions, and atmospheric effects. The primary goal of radiometric correction is to ensure that the pixel values in the imagery accurately represent the true radiance or

reflectance of the objects and features on the Earth's surface (Liang & Wang, 2019). This process is to avoid radiometric errors in the measured reflectance, which includes a correction for the sun's azimuth, elevation, and inconsistencies of radiometric response (gain and bias) between detectors (Pons & Solé-Sugrañes, 1994). It can convert image digital numbers to at-sensor radiance or top-of-atmosphere reflectance.

### *2.2.8 Atmospheric Correction*

When the reflected electromagnetic radiation is recorded by airborne or space-based sensors, the observed radiation may not be the same as the reflected radiation from the same object from a short distance. It is due to the scattering and absorption effects imposed by the ambient that the radiation goes through. Therefore, the recorded spectral response may not correspond to the the plant's biological state but the variation in atmospheric transmittance (Gautam & Pagay, 2020). Atmospheric correction can remove these intervening effects from the atmosphere to estimate the surface reflectance characterizing plant properties.

### *2.2.9 Vegetation Index (VI)*

VI is a term that is usually based on the sum, difference, or ratio of reflectance from two or more spectral bands to derive plants' status (Rouse et al., 1974). These composite indices are more sensitive than the individual band to the plant's biophysical and biochemical features impacted by stresses. The mathematical computation of VI such as normalization and division enables VI to be in comparison by reducing the uncertainties in reflectance resulting from illumination conditions, instrument noise, atmospheric conditions, and soil background (Qi et al., 1995; Van Leeuwen & Huete, 1996). The values derived from the VI are relative, providing a theoretical level ranging from low to high values. Some of them have been demonstrated to be partially correlated with various physiological activities, including crop water status (Zarco-Tejada et al., 2013).

VIs can be classified as chlorophyll and structural VIs, xanthophyll VIs, and water VIs according to their operational principles in plant water status assessment. VIs in the chlorophyll and structural category are related to the concentration of chlorophyll and internal structure in leaves and canopy architecture (Zarco-Tejada et al., 2012), which detect the water deficit indirectly by accounting for the corresponding physiological variances in photosynthesis, greenness, and vigor. However, what they represent is the accumulated influence of water deficit and hence these indices are viewed as long-term responses to water status (Wong & Gamon, 2015). As these responses usually happen with visible symptoms, it is assumed that the capacity of chlorophyll and structural VIs for the evaluation of crop water potential is limited. Transformed Chlorophyll Absorption Reflectance Index, Excess Green Index, Normalized Difference Red Edge Index, and Normalized Difference Vegetation Index are some common VIs being used in this category (Zhang et al., 2021).

Xanthophyll VIs are sensitive to the de-epoxidation state of the xanthophyll cycle and photosynthetic rate, which are linked back to water stress (Gamon et al., 1992). The photochemical reflectance index is a representation of this category. Water VIs detect changes in water content in crops through the establishment of a simple ratio or normalization index between one reflectance within the water absorption bands

and another one outside the water absorption bands used as a reference (Sims & Gamon, 2002). Another combination is one reflectance in NIR and the other one in SWIR. SWIR reflectance reflects the variability of both the water content and the spongy mesophyll structure of crops, and the NIR reflectance is more influenced by the internal structure and dry matter content in leaves. This combination would eliminate the changes caused by internal structure and dry matter content, enhancing the capability of water content retrieval (Wang et al., 2015).

#### *2.2.10 Normalized Difference Spectral Index (NDSI) and Simple Ratio Spectral Index (SRI)*

The VIs mentioned above only make use of a limited number of bands. More information embedded in the hundreds of bands from hyperspectral data might be worth exploring. As normalized difference index and simple ratio index are the most commonly used forms of reflectance indices, they are employed to extract the information hidden in the hyperspectral spectrum. NDSI uses all possible combinations of two spectral bands to determine optimal normalized indices for estimating the target of interest (Inoue et al., 2008; Pôças et al., 2015). SRI was originally established to assess chlorophyll and anthocyanin contents in the VIS domain (Jackson & Huete, 1991). Their performance was evaluated recently, and they were reported as useful approaches in computing meaningful spectral variables (Maimaitiyiming et al., 2017; Pôças et al., 2017; Rapaport et al., 2015; Tosin et al., 2020).

### 2.3 Soil – Atmosphere – Vegetation Interaction

To manage plant water status in vineyards and achieve optimal grape quality, it is necessary to understand the environmental factors that are associated with GWS variability. The water conditions in plants are essentially an integrative response to soil properties (soil texture, water availability, and electrical conductivity in root zone), atmospheric effect (air temperature, rainfall, solar radiation, relative humidity, and wind velocity/ direction), plant characteristics (root depth, genetic control, leaf area, and aerodynamic features), and cultivation practices (Choné, 2001; Green et al., 2005). A study supports the view that cultivar was the main driver of the water status of vines under well-watered conditions, while vegetative expression and soil type became more dominant as the water deficit increased (Taylor et al., 2010). In addition, climatic variables (Irmak & Mutiibwa, 2010) and water table level (Guix-Hébrard et al., 2007) were mentioned as other driving factors characterizing the variation of the plant water budget. These drivers induce hydraulic and metabolic signals that trigger the water consumption by crops, along with a range of physiological responses to water deficit, consequently shaping the spatial and temporal variation of GWS in fields.

Clearly, there is an opportunity, to incorporate these environmental factors as ancillary variables, to provide a robust estimation of GWS. The results from several studies have provided a solid basis that the variables of canopy characteristics, soil properties, and climate factors might serve as a good proxy for the water budget (Acevedo-Opazo et al., 2010; Brillante et al., 2016; Suter et al., 2019; Taylor et al., 2012). This might be able to support GWS estimation across a given vineyard in a relatively affordable manner, while incorporated into irrigation systems to potentially automate the practice. Besides, the important variables and their relationship at different

phenological stages may be determined through this simulation. Moreover, it can assist growers in adopting corresponding cultivation strategies to adjust GWS within different blocks (Taylor et al., 2010).

### *2.3.1 Soil factors*

#### 2.3.1.1 Soil properties

Soil characteristics, particularly the soil moisture content, are directly associated with the soil water availability to crops and stomatal control via hydraulic and abscisic acid (ABA) signals, influencing transpiration when irrigation requirement is not met (Davies & Zhang, 1991; Van Leeuwen et al., 2004). Elevation and slope are some important topographical elements that affect vine development and grape production through adjusting GWS (Jones et al., 2004; Zhang & Shao, 2015). These elements spatially influence which refers to the amount of water that soil can store and release for crop use (Bellvert et al., 2012; Verburg et al., 2017). Plant available water holding capacity (PAWC) is the difference between the amount of water at field capacity (the amount of water soil can hold against gravity) and at permanent wilting point (the amount of water remaining that crops cannot extract anymore). The spatial PAWC determines the soil water availability across the fields, further providing the basis for GWS changes (Van Leeuwen et al., 2009). Soil properties influence GWS based on soil texture, organic matter, and compaction. Fine textured soil increases PAWC because its smaller pores help to hold water against free drainage and further result in higher field capacity. Nevertheless, if the fraction of clay dominates the soil profile, PAWC may decrease due to an increasing permanent wilting point. Compaction can decrease PAWC by reducing the volume of soil pores while squeezing the pores smaller, which adversely influences the PAWC. Organic matter is regarded as increasing PAWC via direct and indirect effects. From a direct point of view, organic matter can enhance field capacity of a given soil even if it is at field capacity. From the indirect perspective, organic matter is beneficial to soil structure and pore stability, which improves the water storage capacity of soil.

Although soil water potential, measured by soil moisture sensors, is a good indicator to evaluate plant water status, the readings usually do not reflect the soil water content within the rhizosphere from vertical and horizontal perspectives. Therefore, a method is required to provide integrative information about the soil moisture profile, which would serve as a better proxy for GWS than single-depth soil moisture readings.

#### 2.3.1.2 Apparent Electrical Conductivity ($EC_a$)

To investigate spatial AWC across the fields, a ground-based sensor method by electromagnetic induction (EMI) technique was proposed by Hezarjaribi and Sourell (2007), because it is fast, affordable, non-destructive, and with high spatial resolution. EM38, with transmitter and receiver induction coils, is one of the most widely used EMI devices in agriculture production (Doolittle & Brevik, 2014). Together with GPS on a mobile platform, continuous geo-referenced point measurements can be collected to generate $EC_a$ maps. Its transmitter coils generate a primary magnetic field inducing an electric current in the soil, which in turn creates a secondary magnetic field and is measured by its receiver coils. The measurement is then converted into an output

voltage that relates linearly to depth-weighted soil $EC_a$ (Rhoades, 1992). The investigated depth range of EM38 can be varied according to the coil orientation (vertical or horizontal) and the distance between the coils (McNeill, 1980). Based on the most common definition, the depth range is up to 1.5 m when using vertical mode with maximum sensitivity at approximately 0.3-0.4 m below the equipment. Under horizontal mode, the depth is up to 0.75 m, and the highest sensitivity is directly below the device.

The outcome of EM38 is a depth-weighted $EC_a$ of a given location which is influenced primarily by clay content and moisture under non-saline conditions (Doolittle & Brevik, 2014; Hedley & Yule, 2009b). It is a function of soil solids, soil solution, and soil water content, which is useful to assess the variability of soil physio-chemical properties (Bittelli, 2011). For example, the EC values of clay are generally between 25 and 100 mS/m, values of sand are between 1 and 10 mS/m, values of rock are lower than 2-3 mS/m, and values of water depend on the dissolved substrates ranging from a few up to 1000 mS/m (Priori, 2017). This information is considered to indirectly estimate soil properties and is relevant to characterize the spatial variation of PAWC (Hedley & Yule, 2009a; Jiang et al., 2007), soil water content (Padhi & Misra, 2011), and GWS (Yu et al., 2020). In addition, the spatial patterns of $EC_a$ are relatively stable under different soil water conditions, with only the difference in intensity (Heil & Schmidhalter, 2017). This information implies that $EC_a$ can be employed to represent the spatial condition of water storage capability and water availability in soil for a considerably large soil volume with high heterogeneity.

### *2.3.2 Meteorological Parameters*

Atmospheric conditions, including relative humidity, solar intensity, air temperature, and wind speed, have a great impact on the variation of GWS by influencing plant transpiration and soil water evaporation. In the context of non-limiting soil water, climatic demands are taken as the main drivers among the environmental variables. Although growers have no control over them, their relationships with plant water status can provide additional information in assessing plant water budgets (Martí et al., 2013; Neukam et al., 2016; Suter et al., 2019).

Plant transpiration rate is mainly controlled by stomatal conductance (gs), and radiation and temperature were identified as the main affecting weather factors due to their influence on vapor pressure deficit (VPD) (Irmak & Mutiibwa, 2010). VPD is a function of air temperature and relative humidity (Nobel, 2009), and decreasing gs is associated with reducing soil water content and high VPD under constant light intensity. Increasing VPD could induce ABA accumulation in the guard-cell apoplast and later result in stomatal closure (Zhang & Outlaw Jr, 2001). Besides, wind speed would decrease aerodynamic resistance and enable the atmosphere to be in contact with the leaf surface, which enhances internal water transport of crop and plant transpiration (Jarvis & McNaughton, 1986). Thus, a drier atmosphere (higher air temperature or lower relative humidity) and higher wind velocity tend to drive higher transpiration rates of plants at the beginning of the events.

### *2.3.3 Plant Factors*

In plant monitoring, the spectral response is influenced to a greater degree by

cultivar, leaf age, leaf thickness, canopy structure/ soil background, and developmental stage (Jones et al., 2009; Mahlein, 2016). When monitoring is conducted at the canopy scale, there are some factors that require extra consideration when using spectral reflectance. Specifically, the leaf area index (LAI, the amount of foliage) and the canopy architecture are critical factors in the outcome of overall canopy reflectance, as they determine the scattering and absorption properties of vegetation. For example, significant absorbance is found in VIS and the latter portion of SWIR regions (1900–2450 nm) no matter whether canopies have high or low LAI (Asner, 1998). Nevertheless, some woody structures such as stems and branches may contribute to the spectral variation in the SWIR region by scattering photons, and hence the detection of water content is varied when the canopy has not reached full closure. For the reflectance from NIR and the former portion of SWIR region (1500–1900 nm), it is a result of the scattering of photons throughout the canopy. In short, the higher the LAI, the larger the intensity of reflectance sensed in these spectra.

#### 2.3.3.1 Scion Characteristics

The size of the canopy coverage is another driver for water consumption by crops. LAI was found to account for 60% of canopy ET during the growing season (Ohana-Levi et al., 2020). A large canopy with a high leaf area intercepts more solar radiation, leading to elevated transpiration through stomata and is thus prone to deplete their water stock (Johnson & Trout, 2012; O'Connell et al., 2010). Thus, larger canopies with more leaf area tend to need more water than smaller canopies. It was also found that the variability in canopy vigor within vineyards can result in grapevines responding differently to the same irrigation regime (Bellvert et al., 2012). During the growing season, canopy size and leaf area are usually modified by cultivation practices involving hedging, trimming, shoot thinning, bud rubbing, and leaf plucking. They affect the plant's microclimate by altering the spatial distribution of leaves and branches, changing the exposure extent of the internal canopy. The temperature within the canopy has been reported to be different from 6 ℃ to 10 ℃ to the ambient during the day (Ferrer et al., 2015), influencing the water consumption level of grapevines.

The inherent characteristic, especially stomatal behavior, plays an important role in water retention and the extent of water deficit by determining the physiological response of plants when encountering water scarcity. Plant species can be classified as isohydric or anisohydric according to the type of stomatal response under dehydration. Isohydric plants close their stomata, also known as active stomatal control, when they detect a drop in soil water content or an increase in atmospheric demand. Anisohydric plants continue to transpire even when soil water content declines, due to the poor stomatal adjustment capacity (Schultz, 2003; Tardieu & Simonneau, 1998). Hence, anisohydric plants tend to maintain photosynthesis at the cost of being damaged by a strong water deficit in comparison with isohydric plants.

Although grapevine is considered tight in stomatal control, some cultivars show more efficient control than others. A study investigated the stomatal behaviors of grapevines through hydric behavior and gas exchange measurement, concluding Sauvignon Blanc as an isohydric variety, Chardonnay as an anisohydric variety, and Pinot Noir as a near-anisohydric variety (Gutiérrez-Gamboa et al., 2019). In the research on grapevines conducted by Levin et al. (2020), the authors showed there was no cultivar difference in stomatal behavior at high and low plant water status,

while the behavior became cultivar-specific during moderate water deficit. The approach of using a pressure bomb to determine GWS seems not to be suitable for isohydric varieties because they maintain a relatively steady leaf water status over a wide range of atmospheric conditions and soil water availability (during the day and during the period of minor-to-moderate water stress) (Limpus, 2009). Pressure bomb can be used on anisohydric varieties to reflect their water stress conditions as stomata is not actively controlled under dehydration (Alvino & Marino, 2017).

#### 2.3.3.2 Rootstock Characteristics

The rootstock cultivars also impose an effect on the drought tolerance of crops. Almost all grapevines are propagated vegetatively and are grafted on Phylloxera-resistant rootstock (Ollat et al., 2016). The phylloxera-resistant rootstocks are derived from American *Vitis* species while the scions with superior fruit quality of the Eurasian species *Vitis vinifera*. As rootstocks make up part of the trunk and the root system, they alter the water status in grafted grapevines in terms of their ability to take up water (Nikolaou et al., 2000). Besides, it has been reported that rootstock impacts vine water relations by affecting scion gas exchange (Soar et al., 2006). The probable mechanisms may include direct hydraulic effect, hormonal control, and chemical signals (Lovisolo et al., 2010).

#### 2.3.3.3 Canopy Age and Size

It is hard to know the effect of age on the drought response of grapevines due to the difficulty in accessing proper plant materials. In the recent research implemented by Nader et al. (2019), the authors compared several physiological parameters of vines planted in 1995, 1971, and 2012. They observed younger vines, compared to older vines, were more vulnerable to dehydration, due to their young and shallower root system preventing vines from accessing deeper water reserves.

## 2.4 Machine Learning Modeling and Statistical Analysis

### *2.4.1 Machine Learning Empirical Models*

In this study, purely empirical methods, also termed “regression” when the response is continuous, are conducted by calibrating a numerical (linear or nonlinear) relationship between one or more predictors (inputs; spectral and environmental variables mentioned in section 2.2 and 2.3) and the response (outputs; $\Psi_{stem}$) (Weiss et al., 2020). There are generally two aims to develop applicable regression models: prediction and inference (James et al., 2021). Prediction is to estimate or predict the response based on a set of predictors, where the exact form of relationships between them is typically not a concern as long as the model can yield an accurate prediction for the response. Inference is to understand which predictors are associated with the response and what is the relationship between them (James et al., 2021). Some advanced ML models and statistical techniques, employed to approach both prediction and inference goals while developing regression models, will be introduced in the following sections.

#### 2.4.1.1 Elastic Net (EN)

EN is a type of regularized linear regression that improves the model fit by limiting the coefficient estimates (toward zero) while reducing the variance of the model (Zou

& Hastie, 2005). Ridge regression and lasso also belong to regularized linear regression, and they both attempt to minimize the sum of squared residuals with some penalty term. The coefficient estimate of each predictor is shrunken towards zero but none of them will become exact zero when using ridge regression. Conversely, lasso not only shrinks the coefficient estimates but also may result in some of them being zero. EN combines the penalty terms from both ridge regression and lasso to keep the advantages in feature selection and dealing with highly correlated variables.

#### 2.4.1.2 Partial Least Squares Regression (PLSR)

PLSR, generally speaking, is the combination of principal component analysis and multiple linear regression (Cho et al., 2007). PLSR identifies a new set of variables that are linear combinations of the original predictors and then fits a linear model via least squares using these new variables. The new variables not only approximate the old predictors but also are related to the response. This model enables dimension reduction while effectively easing the issues of multicollinearity and overfitting. Thus, the use of PLSR in estimating plant moisture content using spectral data is popular (Das et al., 2017; Ji et al., 2007; Ullah et al., 2013).

#### 2.4.1.3 Support Vector Regression (SVR)

SVR is an extension of support vector machine (SVM) for regression purposes (James et al., 2021). SVR attempts to find hyperplanes in n-dimensional dataspaces and define coefficients that minimize the loss, where only residuals larger in absolute value than some positive constants contribute to the loss function. There are several mathematical functions (kernel type) that allow data points to be projected on different planes, and hence SVR provides flexibility during modeling. SVR is considered suitable for handling small datasets (not containing a lot of outliers) without overfitting the model, but this model requires higher memory allocation for execution (Priya & Ramesh, 2020).

#### 2.4.1.4 Random Forest Regression (RFR)

Random forest is an ensemble learner that is based on many weak decision trees, and it is adopted for regression purposes in this study. To apply RFR, the original dataset is divided into in-bag samples for training the trees and out-of-bag samples for internal cross-validation to assess the error during the learning process. Each tree is built independently and parallelly based on the two user-defined attributes, the number of trees, and the number of predictors used to split each node while creating the tree. Subsequently, the average of the predicted response of all the trees makes the prediction. The advantage of RFR is it can perform well on a model that has a non-linear and complex relationship between response and predictors, and the ensemble characteristic enables an average result from individual models for a more reliable outcome. RFR is considered ideal for handling massive datasets because it requires less time in data preprocessing and is easy to implement (Priya & Ramesh, 2020).

#### 2.4.1.5 Multilayer Perceptron (MLP)

MLP is an example of a feedforward artificial neural network that is developed based on the principle of biological neural networks. It is a deep learning model which is the subfield of the ML family. MLP consists of three layers, including the input layer,

hidden layer(s), and output layer, where there are multiple single neurons that perform mathematical transformations to the inputs. The neurons in each layer are fully connected to those in the neighboring layers, and trainable weights are assigned to each connection in the neural network. The characteristics of MLP is that it can approximate complicated nonlinear relationships and capture high-level abstraction by increasing the complexity of model structure (number of layer and neuron) and adjusting the massive numbers of weight to provide flexibility in reducing errors (Priya & Ramesh, 2020; Schmidhuber, 2015). The disadvantage of this model is the need for a considerable time in the training phase.

### *2.4.2 Preprocessing Techniques and Evaluation Methods*

#### 2.4.2.1 Standardization

It is recommended to execute standardization on the dataset beforehand, as most of the ML algorithms and optimization algorithms have better behavior if the predictors are on the same scale. It is computed as follows:

$$x_{i,std} = \frac{x_i - \mu_x}{\sigma_x} \quad (1)$$

Where $x_{i,std}$ is the standardized value of a particular predictor of $i^{th}$ observation, $\mu_x$ is the corresponding mean, and $\sigma_x$ is the corresponding standard deviation.

#### 2.4.2.2 Correlation

It is a statistical measurement that expresses the strength, correlation coefficient (r), of how two variables are related to each other. r ranges between ±1. The closer r is to zero, the weaker the relationship. The positive value of r indicates a positive correlation, which means the values of the two variables tend to increase together. The negative value of r refers to a negative correlation, and the values of two variables tend to move in a contrary way. It can be classified as Pearson correlation (measuring linear correlation) and Spearman correlation (measuring rank correlation based on monotonic relationship).

#### 2.4.2.3 K-fold cross-validation (k-fold CV)

In order to implement model assessment (select the models with better performance) or parameter selection (select the level of flexibility), cross-validation is a method in utilization through computing the estimated test error (James et al., 2021). It is used when there is an absence of a validation set for estimating the error, which commonly happens when the collection of ground truth is labor-intensive and time-consuming. A recommended cross-validation approach is k-fold CV. It randomly splits the original training set into k groups with approximately equal sizes. Each group served as a validation set each time, while the remaining groups made up the training set. Then all the calculated evaluation matrices are used to derive the estimated error. Empirically, k-fold CV performs well with k = 5 or 10. K-fold CV is computed as follows:

$$CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} M_i \quad (2)$$

Where k is the number of groups into which the observation is divided, and $M_i$ is the evaluation matrix computed based on the held out $i^{th}$ group of observations which is

not included in the training dataset.

2.4.2.4 Coefficient of Determination ($R^2$)

$R^2$ ranges between 0 and 1 and indicates the extent to which the responses can be explained by the predictors. $R^2$ near 1 indicates that most of the variance in the response is explained by the regression model, and near 0 indicates the model explains little variability in the response. The $R^2$ is evaluated on the test dataset, another set of observations not used to train the regression model, to select the appropriate regression model while preventing overfitting. Overfitting happens when the regression model performs well on the training data but plays badly on the test data. It is computed as follows:

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

Where n is the number of observations used to fit the model, $y_i$ is the measured value of the response of $i^{th}$ observation, $\bar{y}$ is the mean response values, and $\hat{y}_i$ is the estimated value of the response of $i^{th}$ observation from the regression model.

2.4.2.5 Root Mean Square Error (RMSE)

The RMSE is used to quantify the extent to which the estimated response value for a given sample is close to its measured response value. The value of the RMSE is small if the estimated values are close to the measured values of the response and large if the estimated and measured responses differ substantially. The RMSE is evaluated on the test dataset as well. The RMSE is computed as follows:

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

Where n is the number of observations used to fit the model, $y_i$ is the measured value of the response of $i^{th}$ observation, and $\hat{y}_i$ is the estimated value of the response of $i^{th}$ observation from the regression model.

2.4.2.6 Ratio of Performance to Interquartile Range (RPIQ)

It is defined as the interquartile range of the measured values divided by the RMSE (Luedeling & Luedeling, 2015). As RPIQ accounts for both prediction errors and the variation of major measured values, it is considered to provide a more objective and comparable evaluation of the model performance. The higher the value of RPIQ, the better the prediction capacity of the model.

## 2.5 The Current Status of Global Research about Using Remote Sensing and Geospatial Techniques for GWS Monitoring

The advancements have been made in the recent viticultural water status studies. Beginning with the analysis of relationships between spectrally sensing information and GWS, hyperspectral signatures have been extensively used and devoted to this area for different grape cultivars, along with various statistical processing methods. Rapaport et al. (2015) used PLSR to correlate Cabernet Sauvignon grapevines and $\Psi_{leaf}$ under a controlled environment. It was found the combination of VIS (530–550 nm) and SWIR (around 1500 nm) can be superiorly indicative of the alterations in $\Psi_{leaf}$.

Maimaitiyiming et al. (2017) tested the strength of linear relationships between NDSI and stomatal conductance measured in Chambourcin vineyards in the central United States. It showed the highest correlation happened when using NDSI (reflectance at 603 and 558 nm) and highlighted the importance of the yellow band (570-630 nm) due to its connection with nitrogen and chlorophyll concentration. In north-eastern Portugal, the important hyperspectral bands in simulating $\Psi_{pd}$ of two cultivars (Touriga Nacional and Touriga Franca) were assessed using Spearman correlation and the feature importance from RFR. The selected spectral combination included NDSI (reflectance at 554 and 561 nm), water index (reflectance at 900 and 970 nm), and D1 (the first derivatives of the reflectance at 730 and 706 nm) (Pôças et al., 2017). Rodríguez-Pérez et al. (2018) used hyperspectral reflectance to estimate leaf water content for four grape cultivars (Mencía, Cabernet Sauvignon, Merlot, and Tempranillo) in northern Spain. It identified an important spectral region (1265-1668 nm) when fitting the continuum removal transformed datasets with functional linear regression. Still, in the same region, the hyperspectral information was modeled with $\Psi_{leaf}$ for the Tempranillo cultivar using PLSR. It showed the relevant spectral bands centered on 826 and 1520 nm when fitted with secondary derivative processed data and the spectral band centered on 1450 nm when fitted with raw data (González-Fernández et al., 2019). In a study conducted in Croatia (Zovko et al., 2019), researchers identified specific spectral bands at 1110, 1259, 1264, and 1334 nm. These bands were found to be associated with O-H, C-H, and N-H stretches in water, carbohydrates, and proteins when investigating Babic grapevines using partial least squares-discriminant analysis. Another study conducted in southern France focused on assessing the usability of hyperspectral bands for stem water potential (Ψstem) estimation (Laroche-Pinel et al., 2021). They explored the availability of corresponding multispectral bands from the Sentinel-2 satellite for three different grapevine cultivars: Syrah, Chardonnay, and Grenache. Their analysis revealed that the relevant spectral regions for this estimation task were located in the red edge, NIR, and SWIR parts of the electromagnetic spectrum. Linear regression and ExtraTree regressor were employed for correlation analysis and to highlight the crucial wavelengths. These studies collectively underscore the importance of various spectral regions spanning the VIS, NIR, and SWIR regions of the electromagnetic spectrum. Given this variability, it is imperative to conduct a comprehensive evaluation that considers the combination of all these factors to establish robust hyperspectral data analysis pipelines. This holistic approach becomes particularly relevant when studying Pinot Noir grapevines grown in the Southern Hemisphere.

When applying the correlation between spectral information with GWS using multispectral aerial platforms, the atmospheric vapor hinders the way as it absorbs most of the solar radiation in SWIR, while SWIR is missing in most of the available commercial UAVs (Brook et al., 2020). These facts lead to inadequate accuracy for the estimation of vegetation water content from an aerial perspective. To solve this shortcoming, some researchers have studied the interoperation between datasets sourced differently. Acevedo-Opazo et al. (2010) worked on the extrapolation of $\Psi_{pd}$ from reference locations in the Syrah and Mourvèdre vineyards in southern France. The results highlighted the usefulness of ancillary information (NDVI, apparent soil electrical resistivity, trunk circumference, exposed leaf area, canopy thickness, and weight of pruned wood) in accounting for spatial GWS when fitted with linear

regression ($R^2$ = 0.87). Taylor et al. (2010) sampled $\Psi_{pd}$ for six cultivars based on soil types and vegetation expression to understand the dominant drivers for GWS variation using recursive partitioning analysis in southern France. They identified cultivar had a dominant effect at low stress, and soil type and vegetation expression became dominant as water restriction increased. The effects imposed by soil properties and topography (such as depth, soil water holding capacity, elevation, and $EC_a$) on GWS variability were also mentioned for Pinot Noir in east-northern Spain using k-means clustering (Bellvert et al., 2012), Riesling at east-southern Canada using linear correlation (Willwerth & Reynolds, 2020), and Cabernet Sauvignon at southern South Africa (Jasse et al., 2021). In addition to underground characteristics, the importance of meteorological data and seasonality in influencing GWS were underpinned for Chardonnay in eastern France using gradient boosting machine ($R^2$ = 0.78; RMSE = 85 kPa) (Brillante et al., 2016), for three cultivars (Merlot, Cabernet franc, and Cabernet-Sauvignon) in western France (Suter et al., 2019), and for three cultivars (Petite Sirah, Cabernet Sauvignon, and Merlot) in western United States using RFR ($R^2$ = 0.77; RMSE = 123 kPa) (Tang et al., 2022). However, the previous studies did not collectively consider environmental variables from all perspectives (including temporal, soil/terrain, vegetation, and weather factors) to determine the useful sets of ancillary variables. Moreover, it was found that using all information derived from multispectral aerial images with MLP could improve $\Psi_{stem}$ estimation, either using reflectance ($R^2$ = 0.87; RMSE = 120 kPa) (Poblete et al., 2017) or VIs (r = 0.62) (Romero et al., 2018).

To develop a wide-scale monitoring system, it needs to rely on satellite missions to allow access to data with proper resolution at a regular frequency. The possibility of data interoperation between multispectral images and environmental variables for GWS estimation means that information could be incorporated into a decision-support tool at the satellite scale to improve viticultural irrigation management. The biggest issue with employing satellite images lies in the available spatial resolution. Satellite images with low resolution were found to fail to reliably represent intra-field heterogeneity (Matese et al., 2015). The sample conditions apply to Sentinel-2, especially in the case of small blocks or blocks with complex borders or both, due to the interference of inter-row components (Devaux et al., 2019; Khaliq et al., 2019). This is one of the reasons why the application of satellite images to GWS estimation is often less than satisfactory, such as Pearson's correlation coefficients of 0.32 in the study of Borgogno-Mondino et al. (2018) using Landsat 8 and $R^2$ = 0.40 and RMSE = 260 kPa in the study of Laroche-Pinel et al. (2021) using Sentinel-2. Another issue that should be aware of when employing satellite images is the velocity that the corrected images can be provided after sensing (Devaux et al., 2019). The post-acquisition process includes data transfer and handling by the ground station, along with a combination of atmospheric, radiometric, and geometric corrections. There are often delays in those processes and the subsequent delivery to the end users (Gautam & Pagay, 2020).

**Reference**


Acevedo-Opazo, C., Tisseyre, B., Taylor, J., Ojeda, H., & Guillaume, S. (2010). A model for the spatial prediction of water status in vines (Vitis vinifera L.) using high resolution ancillary information. *Precision Agriculture, 11*(4), 358-378.

Alvino, A., & Marino, S. (2017). Remote sensing for irrigation of horticultural crops. *Horticulturae, 3*(2), 40.

Améglio, T., Archer, P., Cohen, M., Valancogne, C., Daudet, F.-a., Dayau, S., & Cruiziat, P. (1999). Significance and limits in the use of predawn leaf water potential for tree irrigation. *Plant and Soil, 207*(2), 155-167.

Ammoniaci, M., Kartsiotis, S.-P., Perria, R., & Storchi, P. (2021). State of the art of monitoring technologies and data processing for precision viticulture. *Agriculture, 11*(3), 201.

Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment, 64*(3), 234-253.

Baboo, S. S., & Devi, M. R. (2011). Geometric correction in recent high resolution satellite imagery: a case study in Coimbatore, Tamil Nadu. *International Journal of Computer Applications*, *14*(1), 32-37.

Baher, Z. F., Mirza, M., Ghorbanli, M., & Bagher Rezaii, M. (2002). The influence of water stress on plant height, herbal and essential oil yield and composition in Satureja hortensis L. *Flavour and Fragrance Journal, 17*(4), 275-277.

Bellvert, J., Marsal, J., Mata, M., & Girona, J. (2012). Identifying irrigation zones across a 7.5-ha 'Pinot noir'vineyard based on the variability of vine water status and multispectral images. *Irrigation Science, 30*(6), 499-509.

Belward, A. S., & Skøien, J. O. (2015). Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing, 103*, 115-128.

Bittelli, M. (2011). Measuring soil water content: A review. *HortTechnology, 21*(3), 293-300.

Blackburn, G. A. (2007). Wavelet decomposition of hyperspectral data: a novel approach to quantifying pigment concentrations in vegetation. *International Journal of Remote Sensing, 28*(12), 2831-2855. https://doi.org/10.1080/01431160600928625

Borgogno-Mondino, E., Novello, V., Lessio, A., & de Palma, L. (2018). Describing the spatio-temporal variability of vines and soil by satellite-based spectral indices: A case study in Apulia (South Italy). *International Journal of Applied Earth Observation and Geoinformation, 68*, 42-50.

Brillante, L., Mathieu, O., Lévêque, J., & Bois, B. (2016). Ecophysiological modeling of grapevine water stress in burgundy terroirs by a machine-learning approach. *Frontiers in Plant Science, 7*, 796.

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.

Campbell, P. K., Middleton, E. M., McMurtrey, J. E., Corp, L. A., & Chappelle, E. W. (2007, May-Jun). Assessment of vegetation stress using reflectance or fluorescence measurements. *J Environ Qual, 36*(3), 832-845. https://doi.org/10.2134/jeq2005.0396

Cho, M. A., Skidmore, A., Corsi, F., van Wieren, S. E., & Sobhan, I. (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation, 9*(4), 414-424. https://doi.org/10.1016/j.jag.2007.02.001

Choné, X. (2001). Stem Water Potential is a Sensitive Indicator of Grapevine Water Status. *Annals of Botany, 87*(4), 477-483. https://doi.org/10.1006/anbo.2000.1361

Das, B., Sahoo, R. N., Pargal, S., Krishna, G., Verma, R., Chinnusamy, V., Sehgal, V. K., & Gupta, V. K. (2017). Comparison of different uni- and multi-variate techniques for monitoring leaf water status as an indicator of water-deficit stress in wheat through spectroscopy. *Biosystems Engineering, 160*, 69-83. https://doi.org/10.1016/j.biosystemseng.2017.05.007

Datt, B. (1999). Remote sensing of water content in Eucalyptus leaves. *Australian Journal of botany, 47*(6), 909-923.

Davenport, J. R., Stevens, R. G., & Whitley, K. M. (2008). Spatial and temporal distribution of soil moisture in drip-irrigated vineyards. *HortScience, 43*(1), 229-235.

Davies, W. J., & Zhang, J. (1991). Root signals and the regulation of growth and development of plants in drying soil. *Annual review of plant biology, 42*(1), 55-76.

Devaux, N., Crestey, T., Leroux, C., & Tisseyre, B. (2019). Potential of Sentinel-2 satellite images to monitor vine fields grown at a territorial scale. *OENO One, 53*(1), 52-59.

Doolittle, J. A., & Brevik, E. C. (2014). The use of electromagnetic induction techniques

in soils studies. *Geoderma, 223*, 33-45.

Evans, J. R., & Seemann, J. R. (1989). The allocation of protein nitrogen in the photosynthetic apparatus: costs, consequences, and control. *Photosynthesis, 8*, 183-205.

Fernández, J. (2014). Plant-based sensing to monitor water stress: Applicability to commercial orchards. *J Agricultural water management, 142*, 99-109.

Fernández, J. E. (2017). Plant-based methods for irrigation scheduling of woody crops. *J Horticulturae, 3*(2), 35.

Ferrer, M., Echeverría, G., & Gonzalez-Neves, G. (2015). Influence of the microclimate defined by the training system on the vineyard behaviour and the oenological quality of Merlot grapes. *International Journal of Agricultural Sciences and Natural Resources, 2*(4), 95-108.

Fuentes, S., De Bei, R., Collins, M. J., Escalona, J. M., Medrano, H., & Tyerman, S. (2014). Night-time responses to water supply in grapevines (Vitis vinifera L.) under deficit irrigation and partial root-zone drying. *Agricultural Water Management, 138*, 1-9. https://doi.org/10.1016/j.agwat.2014.02.015

Gamon, J., Penuelas, J., & Field, C. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment, 41*(1), 35-44.

Gausman, H. W. (1974). Leaf reflectance of near-infrared. *Photogrammetric Engineering, 40*(2), 183-191.

Gautam, D., & Pagay, V. (2020). A review of current and potential applications of remote sensing to study the water status of horticultural crops. *Agronomy, 10*(1), 140.

Gholamin, R., & Khayatnezhad, M. (2011). The effect of end season drought stress on the chlorophyll content, chlorophyll fluorescence parameters and yield in maize cultivars. *Scientific Research and Essays, 6*(25), 5351-5357.

González-Fernández, A. B., Sanz-Ablanedo, E., Gabella, V. M., García-Fernández, M., & Rodríguez-Pérez, J. R. (2019). Field Spectroscopy: A Non-Destructive Technique for Estimating Water Status in Vineyards. *Agronomy, 9*(8), 427.

Greenspan, M. D., Schultz, H. R., & Matthews, M. A. (1996). Field evaluation of water transport in grape berries during water deficits. *Physiologia Plantarum, 97*(1), 55-62.

Guix-Hébrard, N., Voltz, M., Trambouze, W., Garnier, F., Gaudillere, J.-P., & Lagacherie, P. (2007). Influence of watertable depths on the variation of grapevine water

status at the landscape scale. *European Journal of Agronomy, 27*(2-4), 187-196.

Gutiérrez-Gamboa, G., Pérez-Donoso, A. G., Pou-Mir, A., Acevedo-Opazo, C., & Valdés-Góme, H. (2019). Hydric behaviour and gas exchange in different grapevine varieties (Vitis vinifera L.) from the Maule Valley (Chile). *South African Journal of Enology and Viticulture, 40*(2), 1-1.

Hedley, C., & Yule, I. (2009a). A method for spatial prediction of daily soil water status for precise irrigation scheduling. *Agricultural Water Management, 96*(12), 1737-1745.

Hedley, C., & Yule, I. (2009b). Soil water status mapping and two variable-rate irrigation scenarios. *Precision Agriculture, 10*(4), 342-355.

Heil, K., & Schmidhalter, U. (2017). The application of EM38: Determination of soil parameters, selection of soil sampling points and use in agriculture and archaeology. *Sensors, 17*(11), 2540.

Hezarjaribi, A., & Sourell, H. (2007). Feasibility study of monitoring the total available water content using non-invasive electromagnetic induction-based and electrode-based soil electrical conductivity measurements. *Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage, 56*(1), 53-65.

Huang, Z., Zeng, H., Hamzavi, I., McLean, D. I., & Lui, H. (2001). Rapid near-infrared Raman spectroscopy system for real-time in vivo skin measurements. *Optics letters, 26*(22), 1782-1784.

Hunt Jr, E. R., & Rock, B. N. (1989). Detection of changes in leaf water content using near-and middle-infrared reflectances. *Remote Sensing of Environment, 30*(1), 43-54.

Ihuoma, S. O., & Madramootoo, C. A. (2017). Recent advances in crop water stress detection. *Computers and Electronics in Agriculture, 141*, 267-275. https://doi.org/10.1016/j.compag.2017.07.026

Inoue, Y., Penuelas, J., Miyata, A., & Mano, M. (2008). Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux measurements in rice. *Remote Sensing of Environment, 112*(1), 156-172. https://doi.org/10.1016/j.rse.2007.04.011

Irmak, S., & Mutiibwa, D. (2010). On the dynamics of canopy resistance: Generalized linear estimation and relationships with primary micrometeorological variables. *Water resources research, 46*(8).

Jackson, R. D., & Huete, A. R. (1991). Interpreting vegetation indices. *Preventive*

*veterinary medicine, 11*(3-4), 185.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). Statistical learning. In *An introduction to statistical learning* (pp. 15-57). Springer.

Jarvis, P. G., & McNaughton, K. G. (1986). Stomatal control of transpiration: scaling up from leaf to region. In *Advances in ecological research* (Vol. 15, pp. 1-49). Elsevier.

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, H.-y., Wang, P.-x., & Yan, T.-l. (2007). Estimations of chlorophyll and water contents in live leaf of winter wheat with reflectance spectroscopy. *Guang pu xue yu guang pu fen xi= Guang pu, 27*(3), 514-516.

Jiang, P., Anderson, S. H., Kitchen, N. R., Sudduth, K. A., & Sadler, E. J. (2007). Estimating plant-available water capacity for claypan landscapes using apparent electrical conductivity. *Soil Science Society of America Journal, 71*(6), 1902-1908.

Johnson, L. F., & Trout, T. J. (2012). Satellite NDVI assisted monitoring of vegetable crop evapotranspiration in California's San Joaquin Valley. *Remote Sensing, 4*(2), 439-455.

Jones, G., Snead, N., & Nelson, P. (2004). Geology and wine 8. Modeling viticultural landscapes: A GIS analysis of the terroir potential in the Umpqua Valley of Oregon. *Geoscience Canada, 31*(4), 167-178.

Jones, H., & Schofield, P. (2008). Thermal and other remote sensing of plant stress. *General and Applied Plant Physiology, 34*(1-2), 19-32.

Jones, H. G. (2004, Nov). Irrigation scheduling: advantages and pitfalls of plant-based methods. *J Exp Bot, 55*(407), 2427-2436. https://doi.org/10.1093/jxb/erh213

Jones, H. G. (2012). Remote sensing of plant stresses and its use in irrigation management. VII International Symposium on Irrigation of Horticultural Crops 1038,

Jones, H. G., Serraj, R., Loveys, B. R., Xiong, L., Wheaton, A., & Price, A. H. (2009). Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. *Functional Plant Biology, 36*(11), 978-989.

Jones, H. G., & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Oxford university press.

Kadioglu, A., Terzi, R., Saruhan, N., & Saglam, A. (2012, Jan). Current advances in the investigation of leaf rolling caused by biotic and abiotic stress factors. *Plant Sci, 182*, 42-48. https://doi.org/10.1016/j.plantsci.2011.01.013

Khaliq, A., Comba, L., Biglia, A., Ricauda Aimonino, D., Chiaberge, M., & Gay, P. (2019). Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. *Remote Sensing, 11*(4), 436.

Kozma-Bognár, V., & Berke, J. (2010). New evaluation techniques of hyperspectral data. *Image, 2000*, 18.04.

Laroche-Pinel, E., Albughdadi, M., Duthoit, S., Chéret, V., Rousseau, J., & Clenet, H. (2021). Understanding vine hyperspectral signature through different irrigation plans: A first step to monitor vineyard water status. *Remote Sensing, 13*(3), 536.

Laroche-Pinel, E., Duthoit, S., Albughdadi, M., Costard, A. D., Rousseau, J., Cheret, V., & Clenet, H. (2021). Towards vine water status monitoring on a large scale using sentinel-2 images. *Remote Sensing, 13*(9), 1837.

Lavoie-Lamoureux, A., Sacco, D., Risse, P. A., & Lovisolo, C. (2017, Apr). Factors influencing stomatal conductance in response to water availability in grapevine: a meta-analysis. *Physiol Plant, 159*(4), 468-482. https://doi.org/10.1111/ppl.12530

Levin, A. D., Williams, L. E., & Matthews, M. A. (2020). A continuum of stomatal responses to water deficits among 17 wine grape cultivars (Vitis vinifera). *Functional Plant Biology, 47*(1), 11-25.

Liang, S., & Wang, J. (Eds.). (2019). *Advanced remote sensing: terrestrial information extraction and applications*. Academic Press.

Limpus, S. (2009). Isohydric and anisohydric characterisation of vegetable crops. The classification of vegetables by their physiological responses to water stress.

Loggenberg, K., Strever, A., Greyling, B., & Poona, N. (2018). Modelling water stress in a shiraz vineyard using hyperspectral imaging and machine learning. *Remote Sensing, 10*(2), 202.

Lovisolo, C., Perrone, I., Carra, A., Ferrandino, A., Flexas, J., Medrano, H., & Schubert, A. (2010). Drought-induced changes in development and function of grapevine (Vitis spp.) organs and in their hydraulic and non-hydraulic interactions at the whole-plant level: a physiological and molecular update. *Functional Plant Biology, 37*(2), 98-116.

Luedeling, E., & Luedeling, M. E. (2015). Package 'chillR'.

Mahlein, A. (2016). Present and future trends in plant disease detection. *Plant Disease, 100*(2), 1-11.

Maimaitiyiming, M., Ghulam, A., Bozzolo, A., Wilkins, J. L., & Kwasniewski, M. T. (2017). Early detection of plant physiological responses to different levels of water stress using reflectance spectroscopy. *Remote Sensing, 9*(7), 745.

Martí, P., Gasque, M., & GonzáLez-Altozano, P. (2013). An artificial neural network approach to the estimation of stem water potential from frequency domain reflectometry soil moisture measurements and meteorological data. *Computers and Electronics in Agriculture, 91*, 75-86.

Massonnet, C., Costes, E., Rambal, S., Dreyer, E., & Regnard, J. L. (2007). Stomatal regulation of photosynthesis in apple leaves: evidence for different water-use strategies between two cultivars. *Annals of Botany, 100*(6), 1347-1356.

Matese, A., & Di Gennaro, S. F. (2015). Technology in precision viticulture: A state of the art review. *International journal of wine research, 7*, 69-81.

Matese, A., Toscano, P., Di Gennaro, S., Genesio, L., Vaccari, F., 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*(3), 2971-2990. https://doi.org/10.3390/rs70302971

Matthews, M. A., & Anderson, M. M. (1989). Reproductive development in grape (Vitis vinifera L.): responses to seasonal water deficits. *American Journal of Enology and Viticulture, 40*(1), 52-60.

McNeill, J. (1980). *Electromagnetic terrain conductivity measurement at low induction numbers*. G. Limited.

Meron, M., Grimes, D., Phene, C., & Davis, K. (1987). Pressure chamber procedures for leaf water potential measurements of cotton. *Irrigation Science, 8*(3), 215-222.

Nader, K. B., Stoll, M., Rauhut, D., Patz, C.-D., Jung, R., Loehnertz, O., Schultz, H. R., Hilbert, G., Renaud, C., & Roby, J.-P. (2019). Impact of grapevine age on water status and productivity of Vitis vinifera L. cv. Riesling. *European Journal of Agronomy, 104*, 1-12.

Naor, A., & Cohen, S. (2003). Sensitivity and variability of maximum trunk shrinkage, midday stem water potential, and transpiration rate in response to withholding irrigation from field-grown apple trees. *HortScience, 38*(4), 547-551.

Neukam, D., Böttcher, U., & Kage, H. (2016). Modelling wheat stomatal resistance in hourly time steps from micrometeorological variables and soil water status. *Journal of Agronomy and Crop Science, 202*(3), 174-191.

Nikolaou, N., Koukourikou, M., & Karagiannidis, N. (2000). Effects of various rootstocks on xylem exudates cytokinin content, nutrient uptake and growth patterns of

grapevine Vitis vinifera L. cv. Thompson seedless.

Nobel, P. S. (2009). Chapter 2 - Water. In *Physicochemical and Environmental Plant Physiology (Fourth Edition)* (pp. 44-99). Academic Press. https://doi.org/10.1016/B978-0-12-374143-1.00002-8

O'Connell, M., Whitfield, D., Abuzar, M., Sheffield, K., McClymont, L., & McAllister, A. (2010). Satellite remote sensing of crop water use in perennial horticultural crops. Program and abstracts Australian irrigation conference, Sydney, Australia.

Ohana-Levi, N., Munitz, S., Ben-Gal, A., & Netzer, Y. (2020). Evaluation of within-season grapevine evapotranspiration patterns and drivers using generalized additive models. *Agricultural Water Management, 228*, 105808.

Ollat, N., Peccoux, A., Papura, D., Esmenjaud, D., Marguerit, E., Tandonnet, J., Bordenave, L., Cookson, S., Barrieu, F., & Rossdeutsch, L. (2016). Rootstocks as a component of adaptation to environment. *Grapevine in a changing environment: a molecular and ecophysiological perspective, 1*, 68-108.

Ougham, H. J., Morris, P., & Thomas, H. (2005). 4 The colors of autumn leaves as symptoms of cellular recycling and defenses against environmental stresses. *Current Topics in Developmental Biology*, *66*(1), 135-160.

Padhi, J., & Misra, R. (2011). Sensitivity of EM38 in determining soil water distribution in an irrigated wheat field. *Soil and Tillage Research, 117*, 93-102.

Pastonchi, L., Di Gennaro, S. F., Toscano, P., & Matese, A. (2020). Comparison between satellite and ground data with UAV-based information to analyse vineyard spatio-temporal variability: This article is published in cooperation with the XIIIth International Terroir Congress November 17-18 2020, Adelaide, Australia. Guest editors: Cassandra Collins and Roberta De Bei. *OENO One, 54*(4), 919-934.

Peñuelas, J., Pinol, J., Ogaya, R., & Filella, I. (1997). Estimation of plant water concentration by the reflectance water index WI (R900/R970). *International Journal of Remote Sensing, 18*(13), 2869-2875.

Pinter Jr, P. J., Hatfield, J. L., Schepers, J. S., Barnes, E. M., Moran, M. S., Daughtry, C. S., & Upchurch, D. R. (2003). Remote sensing for crop management. *Photogrammetric Engineering and Remote Sensing, 69*(6), 647-664.

Planet Team. (2017). Planet application program interface: In space for life on earth. *San Francisco*.

Poblete, T., Ortega-Farias, S., Moreno, M. A., & Bardeen, M. (2017, Oct 30). Artificial Neural Network to Predict Vine Water Status Spatial Variability Using

Multispectral Information Obtained from an Unmanned Aerial Vehicle (UAV). *Sensors (Basel), 17*(11). https://doi.org/10.3390/s17112488

Pôças, I., Gonçalves, J., Costa, P. M., Gonçalves, I., Pereira, L. S., & Cunha, M. (2017). Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region. *International Journal of Applied Earth Observation and Geoinformation, 58*, 177-190.

Pôças, I., Rodrigues, A., Gonçalves, S., Costa, P., Gonçalves, I., Pereira, L., & Cunha, M. (2015). Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices. *Remote Sensing, 7*(12), 16460-16479. https://doi.org/10.3390/rs71215835

Pons, X., & Solé-Sugrañes, L. (1994). A simple radiometric correction model to improve automatic mapping of vegetation from multispectral satellite data. *Remote sensing of Environment*, *48*(2), 191-204.

Priori, S. (2017). Sensori ad Induzione Elettromagnetica per il Rilevamento Prossimale del Suolo: Opportunità per L'agricoltura di Precisione e L'individuazione di Anomalie Sottosuperficiali. https://www.georgofili.info/contenuti/sensori-ad-induzione-elettromagnetica-per-il-rilevamento-prossimale-del-suolo-opportunit-per-lagrico/4127

Priya, R., & Ramesh, D. (2020). ML based sustainable precision agriculture: A future generation perspective. *Sustainable Computing: Informatics and Systems, 28*, 100439.

Qi, J., Moran, M., Cabot, F., & Dedieu, G. (1995). Normalization of sun/view angle effects using spectral albedo-based vegetation indices. *Remote Sensing of Environment, 52*(3), 207.

Rallo, G., Minacapilli, M., Ciraolo, G., & Provenzano, G. (2014). Detecting crop water status in mature olive groves using vegetation spectral measurements. *Biosystems Engineering, 128*, 52-68. https://doi.org/10.1016/j.biosystemseng.2014.08.012

Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing, 109*, 88-97. https://doi.org/10.1016/j.isprsjprs.2015.09.003

Rhoades, J. (1992). Instrumental field methods of salinity appraisal. *Advances in measurement of soil physical properties: Bringing theory into practice, 30*, 231-248.

Rienth, M., & Scholasch, T. (2019). State-of-the-art of tools and methods to assess vine

water status. *OENO One, 53*(4), 619-637.

Rodríguez-Pérez, J. R., Ordóñez, C., González-Fernández, A. B., Sanz-Ablanedo, E., Valenciano, J. B., & Marcelo, V. (2018). Leaf water content estimation by functional linear regression of field spectroscopy data. *Biosystems Engineering, 165*, 36-46.

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. https://doi.org/10.1016/j.compag.2018.02.013

Rossini, M., Fava, F., Cogliati, S., Meroni, M., Marchesi, A., Panigada, C., Giardino, C., Busetto, L., Migliavacca, M., Amaducci, S., & Colombo, R. (2013). Assessing canopy PRI from airborne imagery to map water stress in maize. *ISPRS Journal of Photogrammetry and Remote Sensing, 86*, 168-177. https://doi.org/10.1016/j.isprsjprs.2013.10.002

Rouse, J., Haas, R., Schell, J., Deering, D., & Harlan, J. (1974). Monitoring the vernal advancements and retrogradation. *Texas AM Univ, Texas*.

Santos, A. O., & Kaye, O. (2009). Grapevine leaf water potential based upon near infrared spectroscopy. *Scientia Agricola, 66*(3), 287-292.

Satterwhite, M. B., & Henley, J. P. (1990). *Hyperspectral signatures (400 to 2500 nm) of vegetation, minerals, soils, rocks, and cultural features: Laboratory and field measurements*. A. E. T. L. F. B. VA.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks, 61*, 85-117.

Scholander, P. F., Bradstreet, E. D., Hemmingsen, E., & Hammel, H. (1965). Sap pressure in vascular plants: negative hydrostatic pressure can be measured in plants. *Science advances, 148*(3668), 339-346.

Schultz, H. R. (2003). Differences in hydraulic architecture account for near-isohydric and anisohydric behaviour of two field-grown Vitis vinifera L. cultivars during drought. *Plant, cell & environment, 26*(8), 1393-1405.

Shimada, S., Funatsuka, E., Ooda, M., Takyu, M., Fujikawa, T., & Toyoda, H. (2012). Developing the monitoring method for plant water stress using spectral reflectance measurement. *Journal of Arid Land Studies, 22*(1), 251-254.

Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment, 81*(2-3), 337-354.

Sinclair, T., & Ludlow, M. (1985). Who taught plants thermodynamics? The unfulfilled potential of plant water potential. *Functional Plant Biology, 12*(3), 213-217.

Smith, M., & Munoz, G. (2002). Irrigation advisory services for effective water use: a review of experiences. Irrigation Advisory Services and Participatory Extension in Irrigation Management Workshop Organized by FAO-ICID. Montreal, Canada,

Soar, C. J., Dry, P. R., & Loveys, B. (2006). Scion photosynthesis and leaf gas exchange in Vitis vinifera L. cv. Shiraz: mediation of rootstock effects via xylem sap ABA. *Australian Journal of Grape and Wine Research, 12*(2), 82-96.

Sozzi, M., Kayad, A., Marinello, F., Taylor, J., & Tisseyre, B. (2020). Comparing vineyard imagery acquired from Sentinel-2 and Unmanned Aerial Vehicle (UAV) platform. *OENO One, 54*(2), 189-197.

Steele, M. R., Gitelson, A. A., Rundquist, D. C., & Merzlyak, M. N. (2009). Nondestructive estimation of anthocyanin content in grapevine leaves. *American Journal of Enology and Viticulture, 60*(1), 87-92.

Sun, S., Zhang, C., Li, X., Zhou, T., Wang, Y., Wu, P., & Cai, H. (2017). Sensitivity of crop water productivity to the variation of agricultural and climatic factors: A study of Hetao irrigation district, China. *Journal of Cleaner Production, 142*, 2562-2569. https://doi.org/10.1016/j.jclepro.2016.11.020

Suter, B., Triolo, R., Pernet, D., Dai, Z., & Van Leeuwen, C. (2019). Modelling stem water potential by separating the effects of soil water availability and climatic conditions on water status in grapevine (Vitis vinifera L.). *Frontiers in Plant Science, 10*, 1485.

Tang, Z., Jin, Y., Alsina, M. M., McElrone, A. J., Bambach, N., & Kustas, W. P. (2022). Vine water status mapping with multispectral UAV imagery and machine learning. *Irrigation Science*, 1-16.

Tardieu, F., & Simonneau, T. (1998). Variability among species of stomatal control under fluctuating soil water status and evaporative demand: modelling isohydric and anisohydric behaviours. *Journal of experimental botany*, 419-432.

Taylor, J. A., Acevedo-Opazo, C., Pellegrino, A., Ojeda, H., & Tisseyre, B. (2012). Can within-season grapevine predawn leaf water potentials be predicted from meteorological data in non-irrigated Mediterranean vineyards? *OENO One, 46*(3), 221-232.

Taylor, J. A., Acevedo-Opazo, C., Ojeda, H., & Tisseyre, B. (2010). Identification and significance of sources of spatial variation in grapevine water status. *Australian Journal of Grape and Wine Research, 16*(1), 218-226.

Thompson, R., Gallardo, M., Valdez, L., & Fernández, M. (2007). Using plant water

status to define threshold values for irrigation management of vegetable crops using soil moisture sensors. *Agricultural Water Management, 88*(1-3), 147-158.

Tosin, R., Pocas, I., Gonçalves, I., & Cunha, M. (2020). Estimation of grapevine predawn leaf water potential based on hyperspectral reflectance data in Douro wine region. *Vitis, 59*, 9-18.

Toth, C., & Jóźków, G. (2016). Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing, 115*, 22-36.

Ullah, S., Skidmore, A. K., Groen, T. A., & Schlerf, M. (2013). Evaluation of three proposed indices for the retrieval of leaf water content from the mid-wave infrared (2–6μm) spectra. *Agricultural and Forest Meteorology, 171-172*, 65-71. https://doi.org/10.1016/j.agrformet.2012.11.014

Van Leeuwen, C., Friant, P., Chone, X., Tregoat, O., Koundouras, S., & Dubourdieu, D. (2004). Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and Viticulture, 55*(3), 207-217.

Van Leeuwen, C., Trégoat, O., Choné, X., Bois, B., Pernet, D., & Gaudillère, J.-P. (2009). Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? *OENO One, 43*(3), 121-134.

Van Leeuwen, W., & Huete, A. (1996). Effects of standing litter on the biophysical interpretation of plant canopies with spectral indices. *Remote Sensing of Environment, 55*(2), 123-138.

Verburg, K., Cocks, B., Manning, B., Truman, G., & Schwenke, G. (2017). *AP Soil plant available water capacity (PAWC) characterisation of select Liverpool Plain soils and their landscape context*. CSIRO, Australia.

Wang, X., Zhao, C., Guo, N., Li, Y., Jian, S., & Yu, K. (2015). Determining the Canopy Water Stress for Spring Wheat Using Canopy Hyperspectral Reflectance Data in Loess Plateau Semiarid Regions. *Spectroscopy Letters, 48*(7), 492-498. https://doi.org/10.1080/00387010.2014.909495

Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment, 236*, 111402.

Williams, L., & Araujo, F. (2002). Correlations among predawn leaf, midday leaf, and midday stem water potential and their correlations with other measures of soil and plant water status in Vitis vinifera. *Journal of the American Society for Horticultural Science, 127*(3), 448-454.

Willwerth, J. J., & Reynolds, A. G. (2020). Spatial variability in Ontario Riesling Vineyards: I. Soil, vine water status and vine performance. *OENO One, 54*, 311-

333.

Wong, C. Y., & Gamon, J. A. (2015, Apr). Three causes of variation in the photochemical reflectance index (PRI) in evergreen conifers. *New Phytol, 206*(1), 187-195. https://doi.org/10.1111/nph.13159

Woolley, J. T. (1971). Reflectance and transmittance of light by leaves. *Plant Physiology, 47*(5), 656-662.

Yin, W., Zhang, C., Zhu, H., Zhao, Y., & He, Y. (2017). Application of near-infrared hyperspectral imaging to discriminate different geographical origins of Chinese wolfberries. *PLoS One, 12*(7), e0180534. https://doi.org/10.1371/journal.pone.0180534

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*.

Zarco-Tejada, P. J., González-Dugo, V., & Berni, J. A. J. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment, 117*, 322-337. https://doi.org/10.1016/j.rse.2011.10.007

Zarco-Tejada, P. J., González-Dugo, V., Williams, L. E., Suárez, L., Berni, J. A. J., Goldhamer, D., & Fereres, E. (2013). A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sensing of Environment, 138*, 38-50. https://doi.org/10.1016/j.rse.2013.07.024

Zhang, L., Han, W., Niu, Y., Chávez, J. L., Shao, G., & Zhang, H. (2021). Evaluating the sensitivity of water stressed maize chlorophyll and structure based on UAV derived vegetation indices. *Computers and Electronics in Agriculture, 185*, 106174.

Zhang, P., & Shao, M. a. (2015). Spatio-temporal variability of surface soil water content and its influencing factors in a desert area, China. *Hydrological Sciences Journal, 60*(1), 96-110.

Zhang, S. Q., & Outlaw Jr, W. (2001). Abscisic acid introduced into the transpiration stream accumulates in the guard-cell apoplast and causes stomatal closure. *Plant, cell & environment, 24*(10), 1045-1054. Zhou, Y., Lam, H. M., & Zhang, J. (2007). Inhibition of photosynthesis and energy dissipation induced by water and high light stresses in rice. *J Exp Bot, 58*(5), 1207-1217. https://doi.org/10.1093/jxb/erl291

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301-320.

Zovko, M., Žibrat, U., Knapič, M., Kovačić, M. B., & Romić, D. (2019). Hyperspectral remote sensing of grapevine drought stress. *Precision Agriculture, 20*(2), 335-347.

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: | Hsiang-En Wei |
| Name and title of main supervisor: | Dr Miles Grafton, Senior Lecturer |
| In which chapter is the manuscript/published work? | Chapter 3 |
| What percentage of the manuscript/published work was contributed by the student? | 82.5% |

Describe the contribution that the student has made to the manuscript/published work:

The candidate used six hyperspectral data groups, two variable selection methods, and three regression models to process hyperspectral data to evaluate important spectral bands for grapevine water status monitoring.

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

- ( ) **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: | Hsiang-En Wei 數位簽署者：Hsiang-En Wei 日期：2023.04.11 16:45:22 +12'00' | Main supervisor's signature: | Miles Grafton Digitally signed by: Miles Grafton DN: CN = Miles Grafton email = w.grafton@massey.ac.bz C = NZ O = Massey University OU = SAE Date: 2023.05.04 21:18:47 +12'00' |

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

# Chapter 3
# Evaluation of Point Hyperspectral Reflectance and Multivariate Regression Models for Grapevine Water Status Estimation

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

This chapter is dedicated to assessing the first hypothesis – The relationship between spectral data and changes in GWS. This can be revealed using statistical preprocessing techniques and ML algorithms, with an objective of obtaining insights into significant spectral regions relevant to GWS variation, over the spectrum (400-2400 nm). This serves as an essential step in employing spectrally sensed information to monitor GWS. Due to the inherent nature of hyperspectral datasets, including high dimensionality and multicollinearity, there is a need to identify robust data analytics pipelines for downsizing the large dataset into a few GWS-sensitive wavelengths. Various statistical preprocessing techniques and models are evaluated for their capabilities in describing GWS variability, with the determination of statistically related spectral regions.

**Abstract:** Monitoring and management of plant water status over the critical period between flowering and veraison, plays a significant role in producing grapes of premium quality. Hyperspectral spectroscopy has been widely studied in precision farming, including for the prediction of grapevine water status. However, these studies were presented based on various combinations of transformed spectral data, feature selection methods, and regression models. To evaluate the performance of different modeling pipelines for estimating grapevine water status, a study spanning the critical period was carried out in two commercial vineyards at Martinborough, New Zealand. The modeling used six hyperspectral data groups (raw reflectance, first derivative reflectance, second derivative reflectance, continuum removal variables, simple ratio indices, and vegetation indices), two variable selection methods (Spearman correlation and recursive feature elimination based on cross-validation), an ensemble of selected variables, and three regression models (partial least squares regression, random forest regression, and support vector regression). Stem water potential (used as a proxy for vine water status) was measured by a pressure bomb. Hyperspectral reflectance was undertaken by a handheld spectroradiometer. The results show that the best predictive performance was achieved by using simple ratio indices and applying partial least squares regression ($R^2$ = 0.85; RMSE = 110 kPa). Models trained with an ensemble of selected variables comprising multicombination of transformed data and variable selection approaches outperformed those fitted using single combinations. Although larger data sizes are needed for further testing, this study compares 38 modeling pipelines and presents the best combination of procedures for estimating vine water status. This may lead to the provision of rapid estimation of vine water status in a nondestructive manner and highlights the possibility of applying hyperspectral data to precision irrigation in vineyards.

### 3.1 Introduction

Grapevine (Vitis spp.) is considered one of the most important berry crops in the world, due to its commercial derivative—wine. The market price of this product is defined by the quality of harvested berries, and water management applied during the growing season has a significant effect on this quality (Chaves et al., 2010). Inadequate water inputs can harm berry quality as the production of some quality-specific flavor precursors is compromised (Ojeda et al., 2002). Excessive irrigation can result in high vigor and strong vegetative growth, further delaying ripening and generating undesirable flavors in the wine (Van Leeuwen et al., 2009). Hence, maintaining grapevine water status (GWS) within a specific range is critical to quality management, and thus, the growers' profit. Nevertheless, studies have shown vines in a single block exhibit a significant variation of GWS even if they receive the same amount of irrigation (Acevedo-Opazo et al., 2008). This variability becomes more prominent under non-irrigated conditions commonly encountered in viticulture (Ojeda et al., 2005), and this potentially leads to increasing variability in berry composition across the vineyard. Most viticulturists use soil moisture sensors and pressure chambers to characterize the dynamics of GWS throughout different growing stages (Rienth & Scholasch, 2019). These measurements provide viticulturists a reference to help guide management strategies that ensure grape quality. However, soil moisture sensors obtain only localized soil moisture values and often fail to reveal the variability of soil water status at depth and spatially, due to soil heterogeneity (Lavoie-Lamoureux et al., 2017). The pressure chamber, despite providing direct

information regarding GWS, is destructive, labor-intensive, and time-consuming. Sampling surveys by pressure chambers do not accurately show spatial and temporal variation in moisture conditions across vines, making it challenging to use in irrigation scheduling, unless high-density sampling is undertaken (Oumar & Mutanga, 2010). In this context, remote sensing (RS) is a potentially promising method that can be used in a nondestructive and timely manner for GWS optimization.

The theoretical basis of applying RS to assessing GWS is attributed to the interaction between leaf water content and the spectral information contained in visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) regions of the electromagnetic spectrum (Elsayed et al., 2011). In the VIS spectrum, the reflectance response is a cumulative effect of water deficit on the content of leaf pigments and the process of photosynthesis (Sims & Gamon, 2003). In the NIR to SWIR spectrum, a partial response is due to the internal structure of the leaf resulting from reduced water content (Gausman, 1974). The rest of the response originates from four water absorption bands centering at around 970, 1200, 1450, and 1940 nm (Curran, 1989). The reflectance in the SWIR region is also determined by nitrogen and various forms of carbon (i.e., lignin and cellulose) in leaves (Gao & Goetz, 1994; Yin et al., 2017). The spectral signatures, the variation of reflectance by wavelengths, can be received either by multispectral or by hyperspectral sensors. Hyperspectral data, characterized by thousands of bands around 1 nm bandwidth over 350–2500 nm (Rodríguez-Pérez et al., 2018), can provide further insights into the relationship between spectral information and a target parameter of interest. Several successful studies have been reported applying hyperspectral techniques to assessing GWS (Maimaitiyiming et al., 2017; Pôças et al., 2015; Rodríguez-Pérez et al., 2018; Rodríguez-Pérez et al., 2007). To better extract relevant spectral information, it is essential to investigate the full spectrum instead of certain regions (Loggenberg et al., 2018; Zovko et al., 2019).

Nevertheless, when using full-spectrum hyperspectral data, problems related to high dimensionality and multicollinearity occur. These characteristics may violate some assumptions of statistics, for instance, the assumption of independence between variables (Fortin, 2005). Models trained with such data tend to overfit and become less accurate in prediction capability (Li et al., 2014). Overfitting occurs when the regression model learns the training set too well, but generalizes poorly using the test set. These issues also limit the transferability and interpretability of the models, making it difficult to identify the important relationship between predictors and responses. To minimize this disturbance and enhance the sensitivity of hyperspectral data to target indicators, various preprocessing or transformation approaches have been used to address these issues (Vasques et al., 2008). These include using specific bands to form new indices (vegetation indices (Maimaitiyiming et al., 2017)), removing background interference to compare spectral characteristics (continuum removal (Kokaly & Clark, 1999)), or taking the derivative of the reflectance to amplify signals (Zarco-Tejada et al., 2003). In addition, variable selection is a method commonly employed to decrease the complexity and the size of the datasets (Doktor et al., 2014). This process selects a subset of variables that optimally describes the relationship between input data and target indicators. Therefore, subsequent modeling can be improved by avoiding overfitting, and a better generalization is obtained by removing noise and irrelevant variables from the dataset (Chandrashekar & Sahin, 2014). Another efficient way to decrease complexity is feature extraction which creates an independently new set of variables based on the input variables to minimize the issue of dimensionality (James et al., 2013).

Hyperspectral measurement records reflectance at thousands of wavelengths, and each wavelength-based recording (variable) contains only a fragment of the information available in the entire spectrum. Significant information may be lost if just a few variables are utilized. In the study of Romero et al. (2018), the modeling performance was enhanced after taking all variables as inputs instead of using a subset of them. The use of multivariate regression models and ML algorithms showed promise in exploiting the full information contained in hyperspectral data and searching the complex patterns between reflectance and crop water status. These methods include partial least squares regression (Das et al., 2021), random forest regression (Wang et al., 2016), and support vector machines (Axelsson et al., 2013). Moreover, it has been recently reported that the advantages of increasing prediction accuracy of hyperspectral-based studies by combining different methods (Du, Xia, Zhang, et al., 2012). This ensemble approach has been implemented by combining different algorithms or techniques for modeling (Engler et al., 2013; Xu et al., 2019) and for variable selection (Feilhauer et al., 2015; Seijo-Pardo et al., 2019). The performance of the ensemble method was demonstrated to be generally better and more robust than a single method alone.

Some studies have been shown to achieve high accuracies in estimating plant water status, using hyperspectral reflectance (Das et al., 2021; Krishna et al., 2019; Pôças et al., 2017; Rodríguez-Pérez et al., 2018; Tosin et al., 2020). However, few studies have compared the accuracy of modeling performance, based on different, pipelines in terms of multicombination of data transformation methods, variable selection approaches, and multivariate regression models. Besides, the ensemble technique, to our knowledge, has never been tested for its potential in the domain of hyperspectral data—GWS estimation. This study aims to (i) evaluate the performance of various modeling pipelines composed of six transformation data groups, two variable selection approaches, and three multivariate regression models; (ii) examine the modeling performance after applying ensemble techniques in terms of using collective variables from different combinations of transformed data groups and variable selection methods as inputs.

## 3.2 Materials and Methods

### *3.2.1 The Context of the Study Vineyards*

The study vineyards are located at Martinborough in the Greater Wellington Region in New Zealand (NZ). Both vineyards are sited on a complex of young soils overlying gravels, developed from sedimentary alluvium associated with the nearby Ruamahanga and Huangarua Rivers (Figure 3.1). The vineyards are two commercial vineyards owned by Palliser Estate and are named Wharekauhau and Pencarrow. Our study areas in these two vineyards are 6.6 and 6.7 ha, respectively. Chardonnay, Pinot Noir, and Sauvignon Blanc dominate the cultivars in both vineyards. Among them, Pinot Noir is noteworthy for being flavor-rich under controlled water deficit conditions. However, severe water stress is detrimental to the yield. Accordingly, Pinot Noir was chosen as the target cultivar in this study, due to its requirement for relatively precise irrigation management. The Pinot Noir vines were planted in the vineyards in 1998-2000, grafted on rootstock 101-14, and trained with two-cane vertical shoot positioning. Inter- and intra-row planting space is 2.2 × 1.7 m for Wharekauhau and 2.2 × 1.8 m for Pencarrow. The annual growth cycle of grapevine in NZ comprises budburst, shoot growth, and flowering (September–November), fruit set and veraison (December–February) followed by berry development and harvesting

(March–May). Cultivation practices, such as shoot thinning, bud rubbing, and leaf plucking, are regularly conducted from October to December during the growing season. Irrigation is usually not required before flowering. From flowering to veraison, as the management of GWS in this timeframe is the most critical determinant to the final berry quality, irrigation is usually determined using the measurement of a pressure chamber.

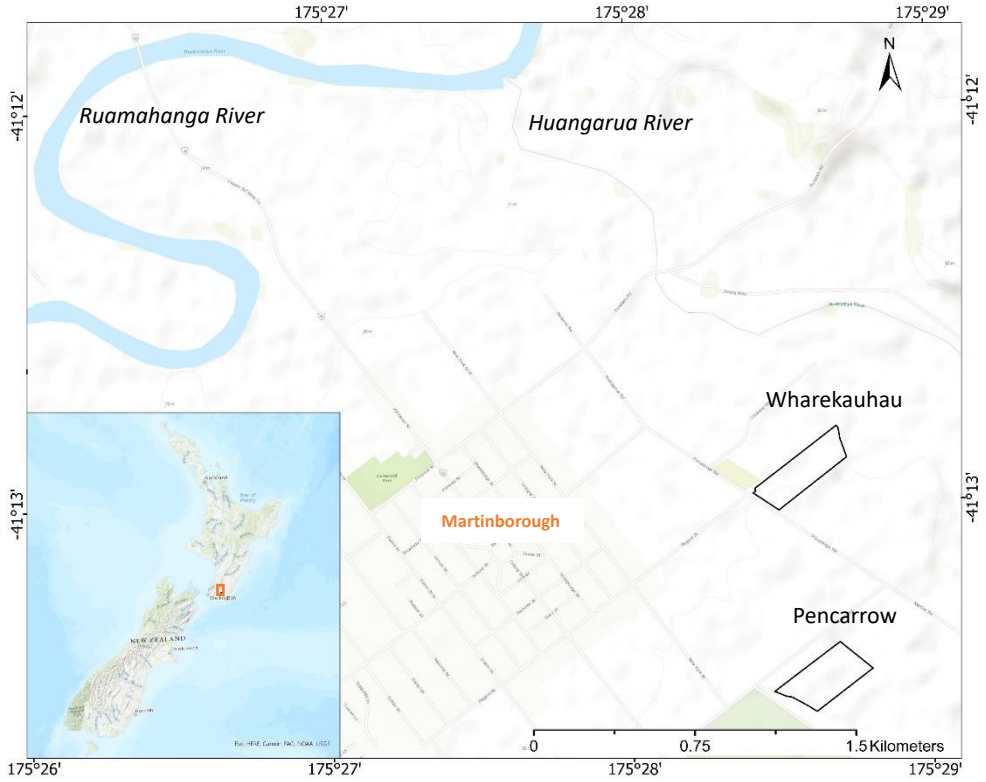



Figure 3. 1 Location of study vineyards.

### *3.2.2 Study Period*

The trials reported in this paper took place from late November 2020 to early February 2021 to match the most critical period for GWS management. The measurement dates, that avoided rain days, were 27 November 2020, 4 December 2020, 14 January 2021, and 22 January 2021 at Wharekauhau, and 4 December 2020, 14 January 2021, 22 January 2021, and 1 February 2021 at Pencarrow.

During the study period, the daily mean temperature varied from 10 to 24 °C, and daily accumulated rainfall ranged between 0 and 30 mm at the vineyards (Figure 3.2). From flowering in late November 2020, several rainfall events occurred in Martinborough, with a maximum daily accumulated rainfall of 30 mm on 10 December. Due to adequate rainfall in late November, the two vineyards were not irrigated in the study period (late November 2020 through to early February 2021) when GWS was a moderate water deficit, desirable for berry quality. At Palliser Estate, the GWS of Pinot Noir during the critical period is expected to keep close to −1300 kPa.

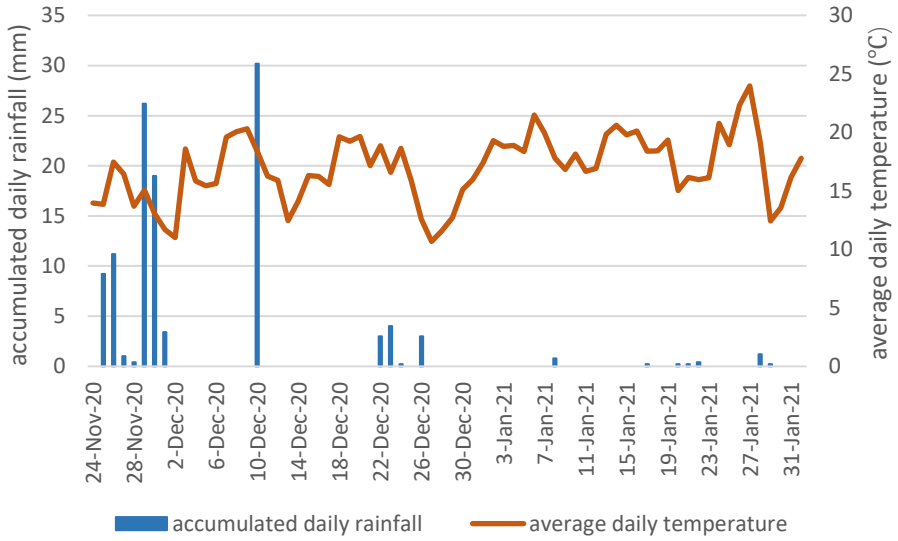



Figure 3. 2 Average daily temperature (red line) and accumulated daily rainfall (blue bars) recorded by the weather station at Palliser Estate between 27 November 2019 and 1 February 2021.

### *3.2.3 Measurement of Vine Stem Water Potential*

Stem water potential ($\Psi_{stem}$) was chosen as a proxy for GWS. As Ψ refers to the suction or the negative pressure, it is usually lower in plants compared to that in soils to enable the absorption of water. The plants naturally maintain a decreasing gradient of Ψ along different parts of the canopy to preserve constant water flow from roots to leaves, later transpiring through the stomata. $\Psi_{stem}$ has been expressed as a comprehensive indicator for early water deficit in vines during the day (Choné, 2001). On each measurement date, several healthy vines were sampled in grids to account for the variability across the vineyards, with two mature and fully expanded leaves from the middle part of the canopy. The mature and fully expanded leaves are more representative of the status of canopies. A pressure chamber model 610 (MPS, Albany, NY, USA) was employed between the hours of 12:00 and 15:00 to assess $\Psi_{stem}$. Prior to the measurement, the sampled leaves were covered with sealable plastic bags for around 1 hour. In this way, transpiration is stopped when the equilibrium of water potential between leaf and stem is attained, which makes this leaf-scale measurement more representative of the canopy conditions. When using the pressure chamber, the pressure is applied onto the scion, which is equal and opposite to the suction in the scion, until the sap is extruded. Therefore, the higher the reading, the more dehydrated the vine is. The two measurements per vine were averaged to represent the canopy water status. A total of 85 separate canopies were surveyed in the two vineyards (Figure 3.3; Table 3.1; Appendix 1).

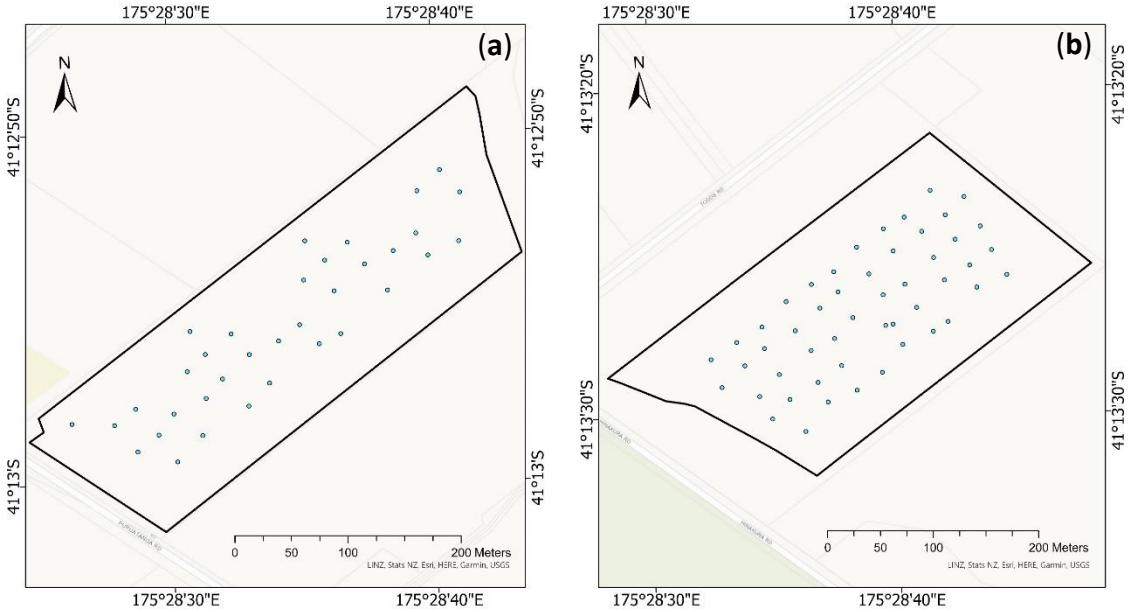



Figure 3. 3 Surveyed canopies (blue points) in Wharekauhau (**a**) and Pencarrow (**b**) vineyards.

Table 3. 1 The number of surveyed canopies on each measurement date.

| | Measurement Data | | | | |
|---|---|---|---|---|---|
| **Vineyard** | **27 November 2020** | **4 December 2020** | **14 January 2021** | **22 January 2021** | **1 February 2021** |
| Wharekauhau | 11 | 8 | 8 | 8 | - |
| Pencarrow | - | 10 | 11 | 11 | 18 |

*3.2.4 Acquisition of Spectral Data and Preprocessing*

Hyperspectral reflectance data were collected by an ASD FieldSpec 4 Hi-Res NG Spectroradiometer (Malvern Panalytical Ltd., Malvern, UK) equipped with a leaf clip and contact probe (touching the surface of the leaf when measuring), providing controlled illumination throughout the field measurements. A white panel ceramic, referencing tile was used for calibrating and referencing the spectrum during the field survey, which was carried out each time before collecting the reflectance data of the next canopy. The reflectance was calculated as the ratio of the optical energy from a sample to the optical energy from the reference panel. The spectral range covers 350–2500 nm with a sampling interval of 1.4 nm between 350–1000 nm and 1.1 nm between 1001–2500 nm. These intervals were then interpolated to 1 nm, resulting in 2151 values for every spectral measurement.

To ensure comparability, the spectral data were obtained during the same timeframe as $\Psi_{stem}$ data measurements. Two leaves per vine, from the same vine used for collecting $\Psi_{stem}$, were selected with similar criteria and positions in the canopies. Measurements were undertaken on the left side and right side of the adaxial surface of each leaf, while avoiding leaf veins, spots, and holes to ensure representative sampling. Five repetitive readings were made at each measuring point, with a total of 20 readings collected per canopy.

Signal instability (noise) was observed at both edges of the electromagnetic spectrum (<400 and >2400 nm), so the reflectance data in these regions were not used for analysis. Each reading was processed using ViewSpec Pro 6.2 software (Analytical Spectral Devices, Inc., Boulder, CO, USA). Splice correction was applied to all the spectra to adjust the mismatches in the visible-near infrared and shortwave infrared two regions. This was achieved by calculating a bias value to help match the shortwave infrared one region at the splice points. These corrected spectra were exported as ASCII text files, and then they were averaged to obtain the mean spectral signatures for the 85 canopies.

### *3.2.5 Data Transformation*

The raw reflectance (the mean hyperspectral signatures of the 85 canopies) were transformed into five feature groups: (i) First derivative, (ii) second derivative, (iii) continuum removal, (iv) simple ratio indices, and (v) vegetation indices.

#### 3.2.5.1 First (1D) and Second (2D) Derivative

Derivative transformations can capture sudden changes over the spectrum and eliminate noise in the baselines (Demetriades-Shah et al., 1990). 1D preprocessing was shown to acquire promising results of modeling leaf water status (Cao et al., 2015; González-Fernández et al., 2019). These transformations were processed using the ViewSpec Pro 6.2 software with a derivative gap of 3, and then exported as ASCII text files, similar to the raw reflectance data. 1D transformation provides the slope of the tangent line of reflectance at a certain wavelength, and 2D indicates the degree to which the slope at a wavelength is changing. There are 2001 variables (corresponding to 2001 wavelengths) in each of the 1D and 2D groups.

#### 3.2.5.2 Continuum Removal (CR)

CR transformation was used to normalize the spectrum to a common baseline. The continuum refers to background absorption. The difference between the measured spectrum and the continuum after transformation was calculated by dividing the raw reflectance values by the corresponding reflectance values of the continuum. This process can highlight absorption characteristics (Kokaly & Clark, 1999), and it is useful for providing other perspectives of hyperspectral signatures other than pure reflectance intensity (González-Fernández et al., 2015). The target bands in this study were determined to be centered at 670, 970, 1175, 1440, and 1925 nm (Table 3.2), due to their direct and indirect relationships to water absorption features (Curran, 1989). This preprocessing was carried out using "FeaturesConvexHullQuotient" from the pysptools library in Python 3.9 to extract several absorption features, including absorption depth, absorption area, continuum slope, width at half maximum of band depth (FWHM), and position of wavelength with minimum reflectance in each of the target bands. This processing produced 25 variables (five target bands × five absorption features per band).

Table 3. 2 Spectral intervals for continuum removal variables.

| Band (nm) | Bandwidth | Central Wavelength (nm) |
|---|---|---|
| 560–750 | 190 | 670 |
| 900–1060 | 160 | 970 |
| 1080–1250 | 170 | 1175 |

| | | |
|---|---|---|
| 1280–1660 | 380 | 1440 |
| 1830–2210 | 380 | 1925 |

3.2.5.3 Simple Ratio Indices (SI)

A study showed that simple ratio indices (SI) using all possible pairwise-band combinations, of reflectance over the entire spectrum, outperformed vegetation indices for predicting the water status of rice (Das et al., 2021). Therefore, in this study, all the possible pairwise-band combinations over 400–2400 nm were used to compute SI (2,001,000 variables in total) using Visual Basic for Applications (VBA) in Excel 2019.

3.2.5.4 Vegetation Indices (VIs)

The most widely used method to extract information from the electromagnetic spectrum is to compute vegetation indices based on the reflectance at certain wavelengths (Pôças et al., 2015; Zarco-Tejada et al., 2013). These indices were designed to enhance spectral features sensitive to target parameters. However, these indices, calibrated based on several databases, utilize only specific regions of the spectrum. It has been reported that they may not be suitable when applied to other datasets (Le Maire et al., 2008). Eleven water status-related vegetation indices in Table 3.3 were calculated for the purpose of comparing $\Psi_{stem}$ estimation fitted with multivariable (raw reflectance, 1D, 2D, CR, and SI) and univariable (VI). The modeling using multivariable as inputs was computed based on multivariable models (partial least square regression, random forest regression, support vector regression), and using univariable as inputs was computed based on linear regression.

Table 3. 3 Vegetation indices used in this study.

| **Vegetation Indices** | **Acronym** | **Formula** | **Reference** |
|---|---|---|---|
| Normalized difference vegetation index | NDVI | (R800−R675)/(R800 + R675) | (Jordan, 1969) |
| Moisture stress index | MSI | R1600/R820 | (Hunt & Rock, 1989) |
| Photochemical reflectance index | PRI | (R531−R570)/(R531 + R570) | (Gamon et al., 1992) |
| Water index | WI | R900/R970 | (Penuelas et al., 1993) |
| Normalized water difference index | NDWI | (R860-R1240)/(R860 + R1240) | (Gao, 1996) |
| Simple ratio water index | SRWI | R860/R1240 | (Zarco-Tejada et al., 2001) |
| Floating position water band index | FWBI | R900/min(R930–980) | (Strachan et al., 2002) |
| Maximum Difference Water Index | MDWI | max(R1500–1750) − min(R1500–1750)/max(R1500–1750) + min(R1500–1750) | (Eitel et al., 2006) |
| Simple ratio index (1300, 1450) | $SI_{1300, 1450}$ | R1300/R1450 | (Seelig et al., 2008) |
| Double difference index | DDI | 2*R1530-R1005-R2055 | (Wang & Li, 2012) |

| | | | |
|---|---|---|---|
| Normalized water balance index | NWBI | (R1500−R538)/(R1500 + R538) | (Rapaport et al., 2015) |

Note: R refers to the reflectance value at a given wavelength.

*3.2.6 Modeling Pipeline*

The total samples (n = 85) were split into training (n = 59) and test (n = 26) sets using a 70/30 ratio. The split was carried out and stratified according to the date of measurement, to ensure that both training and test sets have corresponding percentages, of samples from each date of measurement. The same composition of samples for the training and test sets was used all the way through this study to ensure comparability. Due to the limited size of training sets, validation was implemented for modeling training by applying 10-fold cross-validation to the training set. It randomly splits the training set into k groups, each of approximately equal size. For each recursion, k-1 groups made up the new training set to fit the model, while one group served as the validation set for evaluating performance. Subsequently, the average performance of the algorithm was then calculated. The test dataset was set aside during variable selection and model training and was not used until the evaluation of modeling performance. The splitting process was undertaken using “train_test_split” from the sklearn package in Python 3.9. A total of 38 modeling pipelines were developed for $\Psi_{stem}$ modeling (Table 3.4).

Table 3. 4 The list of modeling pipelines.

| No | Feature Group | Variable Source | Regression Model |
|---|---|---|---|
| 1 | Raw reflectance | Full set | PLSR |
| 2 | 1D reflectance | Full set | PLSR |
| 3 | 2D reflectance | Full set | PLSR |
| 4 | CR variables | Full set | PLSR |
| 5 | SI | Full set | PLSR |
| 6 | Raw reflectance | Full set | RFR |
| 7 | 1D reflectance | Full set | RFR |
| 8 | 2D reflectance | Full set | RFR |
| 9 | CR variables | Full set | RFR |
| 10 | SI | Full set | RFR |
| 11 | Raw reflectance | Spearman correlation-selected variables | RFR |
| 12 | 1D reflectance | Spearman correlation-selected variables | RFR |
| 13 | 2D reflectance | Spearman correlation-selected variables | RFR |
| 14 | CR variables | Spearman correlation-selected variables | RFR |
| 15 | SI | Spearman correlation-selected variables | RFR |
| 16 | Raw reflectance | RFECV-selected variables | RFR |
| 17 | 1D reflectance | RFECV-selected variables | RFR |
| 18 | 2D reflectance | RFECV-selected variables | RFR |
| 19 | CR variables | RFECV-selected variables | RFR |
| 20 | SI | RFECV-selected variables | RFR |
| 21 | Raw reflectance | Full set | SVR |
| 22 | 1D reflectance | Full set | SVR |
| 23 | 2D reflectance | Full set | SVR |
| 24 | CR variables | Full set | SVR |
| 25 | SI | Full set | SVR |
| 26 | Raw reflectance | Spearman correlation-selected variables | SVR |
| 27 | 1D reflectance | Spearman correlation-selected variables | SVR |
| 28 | 2D reflectance | Spearman correlation-selected variables | SVR |
| 29 | CR variables | Spearman correlation-selected variables | SVR |
| 30 | SI | Spearman correlation-selected variables | SVR |
| 31 | Raw reflectance | RFECV-selected variables | SVR |
| 32 | 1D reflectance | RFECV-selected variables | SVR |
| 33 | 2D reflectance | RFECV-selected variables | SVR |
| 34 | CR variables | RFECV-selected variables | SVR |

| | | | |
|---|---|---|---|
| 35 | SI | RFECV-selected variables | SVR |
| 36 | - | Ensemble of selected variables | RFR |
| 37 | - | Ensemble of selected variables | SVR |
| 38 | VI | Single variable | LR |

Note: “No” refers to the assigned number of each pipeline, “1D” refers to the first derivative, “2D” refers to the second derivative, “CR” refers to continuum removal, “SI” refers to simple ratio indices, “VI” refers to vegetation indices, “RFECV” refers to recursive feature elimination based on cross-validation, “PLSR” refers to partial least squares regression, “RFR” refers to random forest regression, “SVR” refers to support vector regression, and “LR” refers to linear regression. Full set refers to the utilization of all variables without being selected.

### *3.2.7 Variable Selection*

Since hyperspectral information is a high dimensional dataset, variable selection assists in reducing the number of variables to the most significant ones, preventing overfitting and improve the prediction performance of the regression models (Guyon & Elisseeff, 2003). In this study, Spearman correlation and recursive feature elimination based on cross-validation were chosen for variable selection for the five feature groups (raw reflectance, 1D, 2D, CR, and SI).

#### 3.2.7.1 Spearman Correlation

Spearman correlation was used to determine the strength and direction of the monotonic relationship between ranked response (the $\Psi_{stem}$ of each vine) and ranked predictor variables (the spectral data at each wavelength). With this monotonic relationship, the paired variables tend to change together, but not necessarily at a constant rate. This method was used to detect nonlinear relationships, and there is no requirement for the variables to be normally distributed. The Spearman correlation coefficient varies between +1 and −1. The closer to ±1, the stronger the monotonic relationship. Variables with coefficients higher than 0.6 were selected in this study. This correlation was implemented using “spearmanr” from the scipy library in Python 3.9.

#### 3.2.7.2 Recursive Feature Elimination Based on Cross-Validation (RFECV)

RFECV performs variable elimination by repetitively constructing the wrapped model and identifying the least ranked variable after each iteration. The least ranked variable is then discarded, and the model is reconstructed using the remaining variables. For SI, 1% of total variables instead of the least one ranked variable was removed at each iteration in this study, due to computational capacity. The process is recursively repeated on a smaller and smaller set of variables until a specified criterion has been reached. RFECV was employed, due to its effect of reducing correlation between predictor variables (Gregorutti et al., 2017), and, to our knowledge, this method has not been investigated for GWS estimation using hyperspectral data. In this study, the criterion was set to use 10-fold cross-validation to automatically determine the best number of variables according to the value of the coefficient of determination ($R^2$). This step was implemented using “RFECV” from the sklearn library in Python 3.9. Random forest regression and linear support vector regression were used as wrapped algorithms to rank variables based on their attributes of feature importance and coefficient, respectively.

3.2.7.3 The Ensemble of Selected Variables

An ensemble of multimethods improved the result of modeling by overcoming the potential problem of a single technique (Du, Xia, Chanussot, et al., 2012). In this study, various subsets of variables were selected by Spearman correlation and RFECV for each feature group, further feeding the models to compare the estimation accuracy. The variables in every feature group that was selected as inputs for random forest regression and support vector regression, with the best performance, were merged to form the ensemble of selected variables. This new set of variables was used to fit the regression models and then be evaluated for their effect on modeling performance.

*3.2.8 Regression Models*

Partial least squares regression (PLSR), random forest regression (RFR), and support vector regression (SVR) were applied to estimate $\Psi_{stem}$ based on hyperspectral reflectance. They were implemented using "PLSRegression", "RandomForestRegressor", and "SVR" from the sklearn library in Python 3.9, respectively. As the performance of regression models is influenced by their parameters (also called hyperparameters), it is necessary to tune the hyperparameters beforehand to prevent overfitting. This enables the regression algorithms to exploit their potential. Grid searching with 10-fold cross-validation, based on the $R^2$ value, was used to search for the best combination of hyperparameters. A list of tuned parameters and their ranges for each algorithm is displayed in Table 3.5. The combination of hyperparameters contributing to the models with the highest $R^2$ values was considered as optimized. These parameters would then be used for later evaluation of model performance on the test set. This technique was carried out using "GridSearchCV" from the sklearn library in Python 3.9 (Appendix 4). For PLSR, the variable extraction processing goes along with modeling, so PLSR used the transformed datasets directly without carrying out any variable selection beforehand. For RFR and SVR, they were trained either with the full set of variables or with selected variables.

Table 3. 5 The tuned hyperparameters and their ranges for each regression model.

| **Regression Model** | **Hyperparameter** | **Range** |
|---|---|---|
| Partial least squares regression | Number of components | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 |
| Random forest regression | The number of variables to be considered for the best split | "auto", "sqrt", "log2" |
| | The maximum depth of the tree | 1 or 2 |
| | The number of trees in the forest | 500 |
| Support vector regression | The used kernel type | "linear", "poly", "rbf", "sigmoid" |
| | Kernel coefficient | "scale", "auto" |
| | Regularization parameter | 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50, 100 |
| | The width of the epsilon-tube | 0.1, 0.3, 0.5, 0.7, 0.9 |

Notes: "Auto" refers to the total number of variables, "sqrt" refers to the squares root of the total number of variables, "log2" refers to the binary logarithm of the

total number of variables, “poly” refers to polynomial, “rbf” refers to radial basis function, “scale” refers to the use of 1/(total number of the variable × variance of the variables) as the kernel coefficient, and “auto” refers to the use of 1/(total number of variable) as the kernel coefficient.

#### 3.2.8.1 Partial Least Squares Regression (PLSR)

PLSR carries out dimensional reduction through generating independent components which linearly integrate the maximum variance in the predictor variables under the supervision of response variables (Martens & Naes, 1992). It then performs least squares regression on these components with the response variables. This technique is useful in addressing datasets with problems associated with multicollinearity and high dimensionality, and preventing overfitting (Feilhauer et al., 2015). Both predictor and response variables were scaled during modeling. PLSR evaluates the significance of each variable by calculating the variable importance of projection (VIP) (Eriksson et al., 2013). The higher the VIP value of a variable, the more important the corresponding spectral data is to the PLSR.

#### 3.2.8.2 Random Forest Regression (RFR)

RFR is an ensemble learning algorithm that contains a large set of regression trees (Breiman, 2001). It uses different bagged samples (from the training set with replacement) to fit those regression trees, and at each node, the trees perform binary splitting using a subset of the input variables. The variable determined for splitting is based on the degree of reduction in the residual sum of squares. The final predicted value of a sample is computed by averaging the prediction of all regression trees. RFR can be used to model nonlinear relationships between variables. It performs well when building on a limited number of samples with a large number of variables, and it has been observed in literature to be robust despite the introduction of noise and bias to the data (Doktor et al., 2014).

#### 3.2.8.3 Support Vector Regression (SVR)

SVR is an extension of the support vector machines specifically designed for regression problems (Smola & Schölkopf, 2004). It calculates a hyperplane in multidimensional space that encompasses the maximum number of samples within the decision boundary lines. It contains kernel functions that transform input space, to required high-dimensional space, and is, thus, able to deal with nonlinearity. Support vector machines processing has proven to be robust when addressing high dimensional datasets for classification problems (Hua et al., 2005). It is less prone to overfitting and has a relatively high generalized performance (Axelsson et al., 2013), even with a limited number of samples (Stamenkovic et al., 2013). In this study, all predictor variables were standardized to have the same scale before SVR processing.

### *3.2.9 Modeling Performance Evaluation*

To compare the performance of regression models for evaluating the impacts of data transformation techniques, variable selection approaches, and regression models, the coefficient of determination ($R^2$) and root mean square error (RMSE) values were computed by applying the trained models with optimized hyperparameters on the test set.

3.2.9.1 Coefficient of Determination ($R^2$)

$R^2$ values range between 0 and 1 to indicate the extent to which the responses can be explained by the predictors. An $R^2$ value near 1 indicates that most of the variance in the response variables is explained by the model, and values nearer 0 indicate that the model explains little of the variance in the responses. The $R^2$ value was computed as follows:

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

where *n* is the number of samples used to fit the model, $y_i$ is the measured value of response of the *i*th sample, $\bar{y}$ is the mean response value, and $\hat{y}_i$ is the estimated value of response of the *i*th sample from the regression model.

3.2.9.2 Root Mean Square Error (RMSE)

The RMSE was used to quantify the extent to which the estimated response value for a given sample matches its measured response value. The value of the RMSE is small if the estimated values are close to the measured values of responses and are large if the estimated and measured responses differ substantially. The RMSE was computed as follows:

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

where *n* is the number of samples used to fit the model, $y_i$ is the measured value of the response of the *i*th sample, and $\hat{y}_i$ is the estimated value of the response of the *i*th sample using the regression model.

## 3.3 Results

### *3.3.1 Variation in Stem Water Potential*

Each vineyard was visited four times, from flowering through to veraison until netting the vines to protect from birds. This timeframe is the most critical in terms of the effects of GWS on berry quality before harvest. The underlying premise is that precise monitoring of GWS in this period is essential to produce grapes with premium quality. Figure 3.4 displays the variation of stem water potential ($\Psi_{stem}$) collected from different canopies over the five field survey days, and the distribution of the measurements on each date. There were no irrigation events during the field surveys, due to the adequate rainfall from late November to early December (Figure 3.4). As the recorded $\Psi_{stem}$ is equal and opposite to the GWS, the effect of heavy rainfall was reflected by the high GWS observed at the initial survey. The subsequently increasing $\Psi_{stem}$ values resulted from no irrigation plus little precipitation indicate the impact of water deficit gradually accumulating in the canopy as the survey proceeded from flowering to veraison. In terms of the full dataset, the collected leaf samples show a range of $\Psi_{stem}$ values from 310 to 1344 kPa (Table 3.6). The frequency histogram in Figure 3.5 reveals the distribution of collected $\Psi_{stem}$ values, and the class number was determined by the square root of the total sample number.

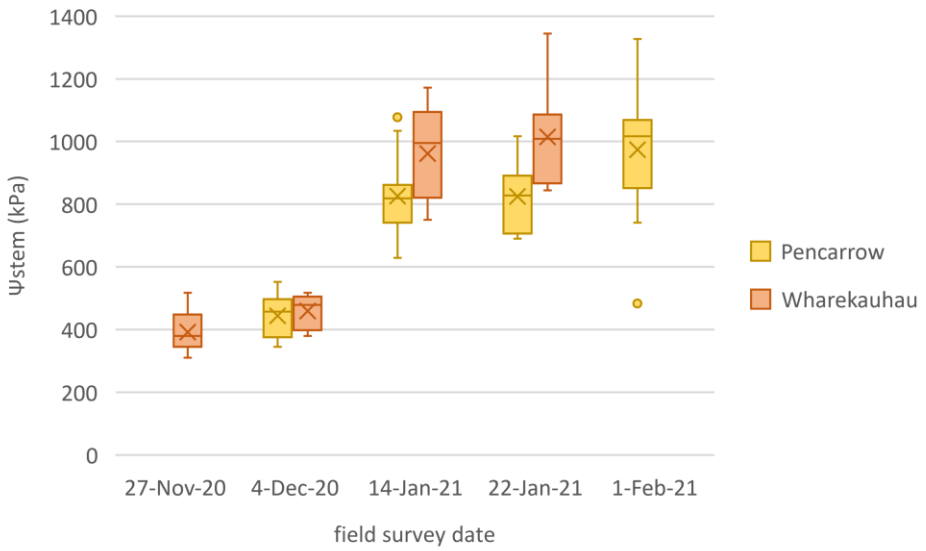



Figure 3. 4 Boxplot of stem water potential ($\Psi_{stem}$) values for the full set of samples collected at Wharekauhau (n = 35) and Pencarrow (n = 50) vineyard.

Table 3. 6 Descriptive statistics of $\Psi_{stem}$ (kPa) of all the observations (n = 85).

| | Total Sample Size | Mean | Standard Deviation | Maximum | Minimum |
|---|---|---|---|---|---|
| $\Psi_{stem}$ (kPa) | 85 | 752 | 277 | 1344 | 310 |

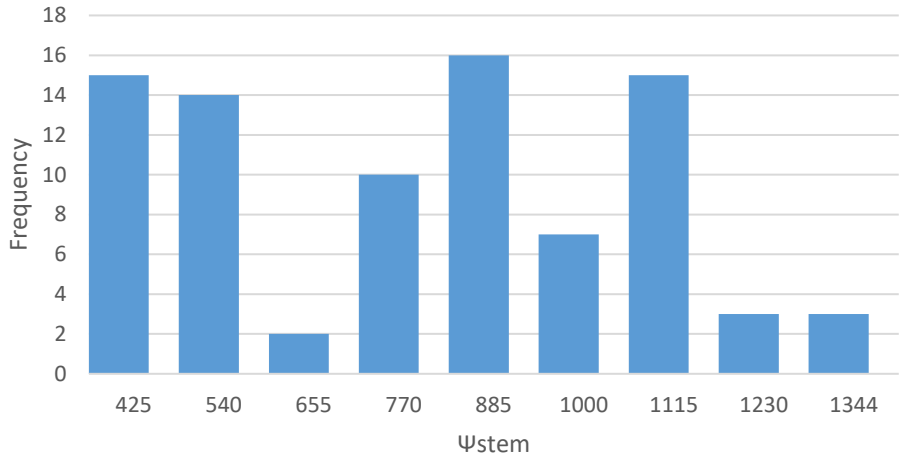



Figure 3. 5 Distribution of $\Psi_{stem}$ (kPa) for all the samples (n = 85). The dataset is not normally distributed.

*3.3.2 Variation in Hyperspectral Data*

The raw reflectance in Figure 3.6a portrays typical reflectance patterns of healthy vegetation: Moderate reflectance at around 500–600 nm, due to the reflection of green light, strong reflectance at around 750–1300 nm, due to the healthy internal structure of leaves, two weak water absorption regions at around 970 and 1200 nm (NIR region), as well as two strong water absorption regions at around 1450 and 1900 nm (SWIR region). Reflectance differences at specific wavelengths over the full spectrum between samples will potentially enable us to estimate GWS for each observation. The spectral regions with evident dispersion of derivative curves may

also be potentially linked to the vine's hydration state, involving 400–800, 1300–1500, 1700, and 1900 nm for the first derivative reflectance, as well as 400–600, 700–1300, 1500–1900, and 2000–2400 nm for the second derivative reflectance (Figure 3.6b,c).

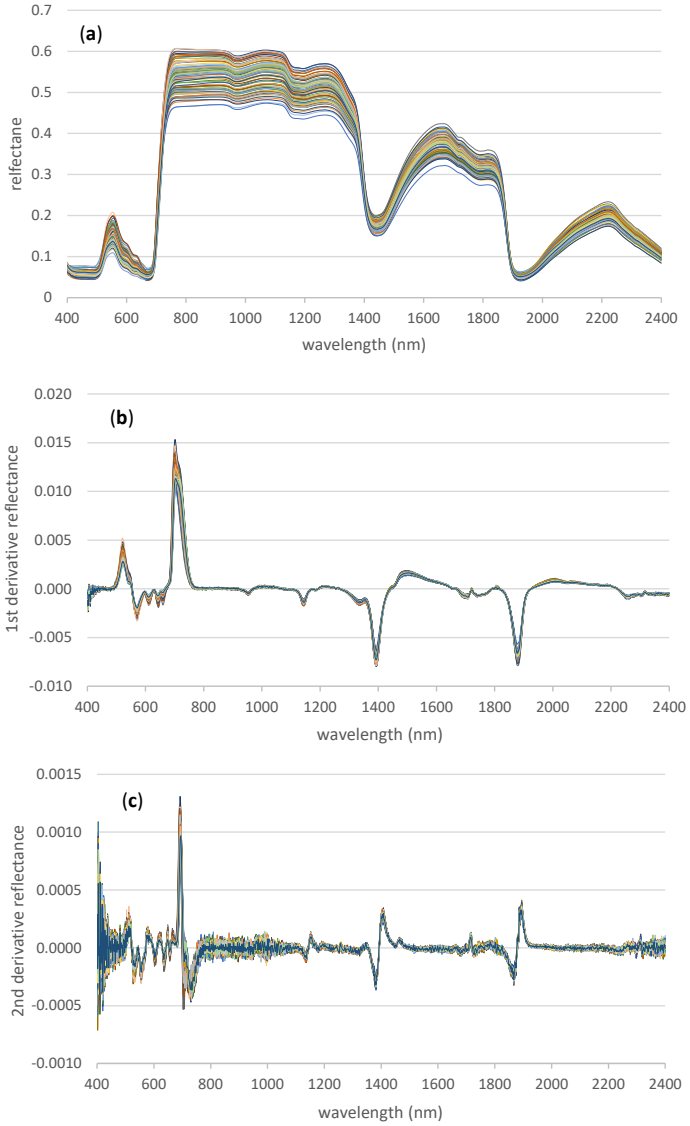



Figure 3. 6 Raw hyperspectral signatures (**a**) and their first (**b**) and second (**c**) derivatives for all samples (n = 85).

### *3.3.3 Modeling Performance*

This study used combinations of six feature groups, two variable selection methods, an ensemble of selected variables, and three regression models to construct 38 modeling pipelines (Table 3.4). Specifically, feature groups include raw reflectance, first derivative (1D) reflectance, second derivative (2D) reflectance, continuum removal (CR) variables, simple ratio indices (SI), and vegetation indices (VI). Variable selection methods include Spearman correlation and recursive feature elimination based on cross-validation (RFECV). Regression models include partial least squares regression (PLSR), random forest regression (RFR), and support vector regression (SVR). The modeling pipelines can be viewed as the pipelines without a variable selection component (i.e., pipelines with numbers 1–5, and 38 in Table 3.4), and with variable selection components (i.e., pipelines with numbers 6–37 in Table 3.4). The

model evaluation metrics (i.e., $R^2$ and RMSE) were computed for modeling performance evaluation. For pipelines with numbers 6–37 in Table 3.4, only pipelines with the best performance on the test set are presented (Table 3.7). The results implied by $R^2$ are equivalent to those implied by RMSE, as expected, since the best performance of modeling is based on the highest value of $R^2$ along with the lowest value of RMSE. Amongst all the modeling pipelines (i.e., pipelines with numbers 1–38 in Table 3.4), the best performance occurs when PLSR was trained with SI, resulting in the highest $R^2$ (0.85) and lowest RMSE (110 kPa). Its scatter plot, shown in Figure 3.7, presents the relationship between observed and predicted $\Psi_{stem}$. Either Spearman correlation or RFECV results improve the performance of RFR and SVR for all feature groups, except for CR variables. Amongst pipelines with numbers 6–37 in Table 3.4, SVR trained with the ensemble of selected variables results in the best performance. The modeling performance of $\Psi_{stem}$ by VIs is poor, as none of the VI resulted in modeling with $R^2$ higher than 0.5. The best performance among all the models using VI as input variable was the one regressed with photochemical reflectance index ($R^2$ = 0.41; RMSE = 210 kPa) (Appendix 6).

Table 3. 7 Results of modeling performance on the test dataset. $R^2$ refers to coefficient of determination. RMSE refers to root mean square error. RPIQ refers to ratio of performance to interquartile range. Full set refers to the utilization of all variables without being selected. RFECV refers to recursive feature elimination based on cross-validation.

| | **Metric** | **Partial Least Squares Regression** | **Random Forest Regression** | **Support Vector Regression** |
|---|---|---|---|---|
| Feature group | | | | |
| Raw reflectance | $R^2$ | 0.81 | 0.70 | 0.74 |
| | RMSE | 123 | 152 | 141 |
| | RPIQ | 3.83 | 3.1 | 3.34 |
| | Variable source | Full set | RFECV | RFECV |
| First derivative reflectance | $R^2$ | 0.79 | 0.70 | 0.67 |
| | RMSE | 127 | 154 | 161 |
| | RPIQ | 3.71 | 3.06 | 2.93 |
| | Variable source | Full set | Spearman correlation | Spearman correlation |
| Second derivative reflectance | $R^2$ | 0.65 | 0.71 | 0.68 |
| | RMSE | 166 | 150 | 158 |
| | RPIQ | 2.84 | 3.14 | 2.98 |
| | Variable source | Full set | Spearman correlation | Spearman correlation |
| Continuum removal variables | $R^2$ | 0.70 | 0.66 | 0.63 |
| | RMSE | 152 | 162 | 170 |
| | RPIQ | 3.1 | 2.91 | 2.77 |
| | Variable source | Full set | Full set | Full set |
| Simple ratio indices | $R^2$ | 0.85 | 0.67 | 0.78 |
| | RMSE | 110 | 160 | 131 |
| | RPIQ | 4.28 | 2.94 | 3.6 |
| | Variable source | Full set | RFECV | RFECV |
| N/A | $R^2$ | N/A | 0.68 | 0.79 |
| | RMSE | N/A | 159 | 128 |
| | RPIQ | N/A | 2.96 | 3.68 |
| | Variable source | N/A | Ensemble of selected variables | Ensemble of selected variables |

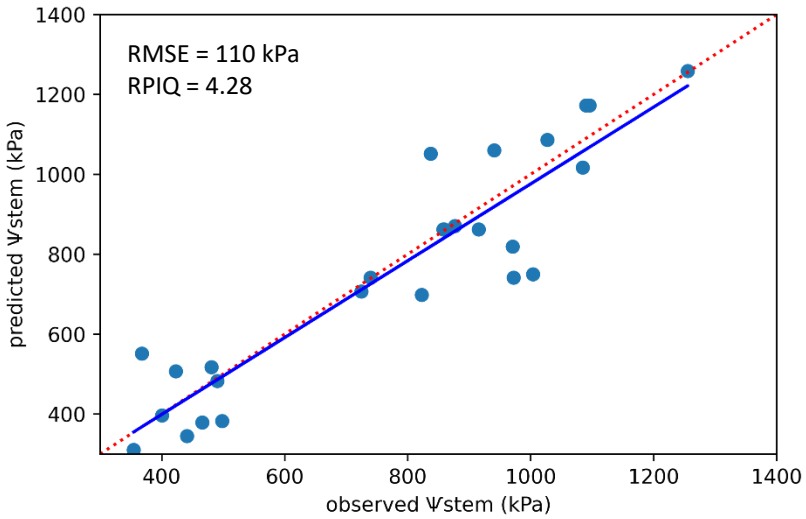



Figure 3. 7 Scatter plots between predicted and observed $\Psi_{stem}$ (kPa) simulated on the test set (n = 26) using the top-performing models—partial least squares regression (PLSR) trained with simple ratio indices (SI). The blue line is the regression line, and the red dotted line is the 1:1 line.

### *3.3.4 Selected Variables and Their Relative Importance*

Figures 3.8–3.12 present the variable importance for each of the five feature groups (raw reflectance, 1D, 2D, CR, and SI). PLSR uses variable importance in projection (VIP), RFR uses variable importance, and SVR uses coefficient to rank the significance for either the full set of variables, Spearman correlation-selected variables, or RFECV-selected variables. For pipelines with numbers 6–37 in Table 3.4, the variable importance was only calculated and presented for the variable subset used as inputs for the modeling pipelines with the best performance on the test set. That is, using SVR to compute variable importance for RFECV-selected variables from raw reflectance, using RFR to compute variable importance for Spearman correlation-selected variables from 1D reflectance, using RFR to compute variable importance for Spearman correlation-selected variables from 2D reflectance, using RFR to compute variable importance for the full set of variables of CR variables, and using SVR to compute variable importance for RFECV-selected variables from SI. This computation was based on the training set, and it can help elucidate significant regions and wavelengths relevant to $\Psi_{stem}$ estimation.

#### 3.3.4.1 Raw Reflectance

When using raw reflectance data, PLSR-based variable importance seems to be distributed evenly across 400–2400 nm. Although PLSR does not conduct variable selection, it calculates VIP for each wavelength to account for their significance in model construction. VIP values higher than one are generally considered significant. The VIP values for raw reflectance data fluctuate around one, with maxima occurring at around 400, 520–630, 700, 1890, and 2400 nm (Figure 3.8a). Variables selected by REFCV disperse at around 400–430, 720, 1049, 1400, 1565–1595, 1890, 2250, and 2370 nm (Figure 3.8b). Their variable importance computed by SVR indicates the region around 400–430 is the most important.

(**a**) (**b**)

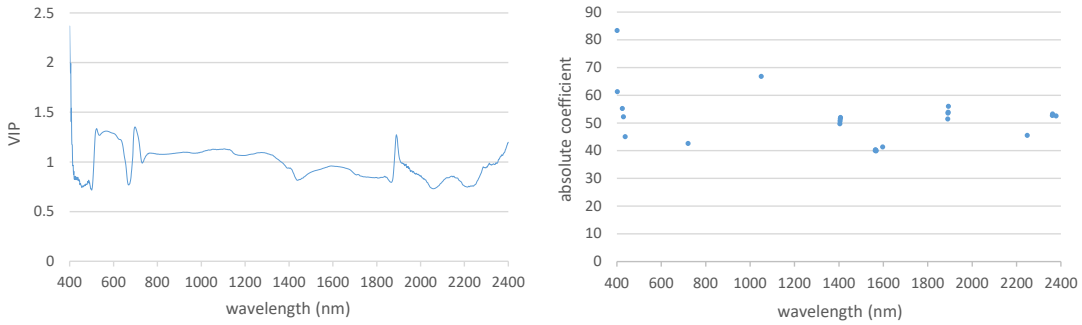



Figure 3. 8 Variable importance for raw reflectance data computed by PLSR (**a**), and support vector regression (SVR) based on variables selected by cross-validated recursive feature elimination (RFECV) (**b**).

3.3.4.2 First Derivative

The VIP values derived from PLSR for each 1 D variable are more discrete compared to that for each raw reflectance variable, with the difference between the highest and lowest VIP values being larger (Figure 3.9a). With Spearman correlation, selected variables concentrate at around 400, 715–760, 800–1250, 1000–1870, and 2250–2350 nm (Figure 3.9b). The important regions computed by RFR are at around 740, 1220, and 1700 nm.

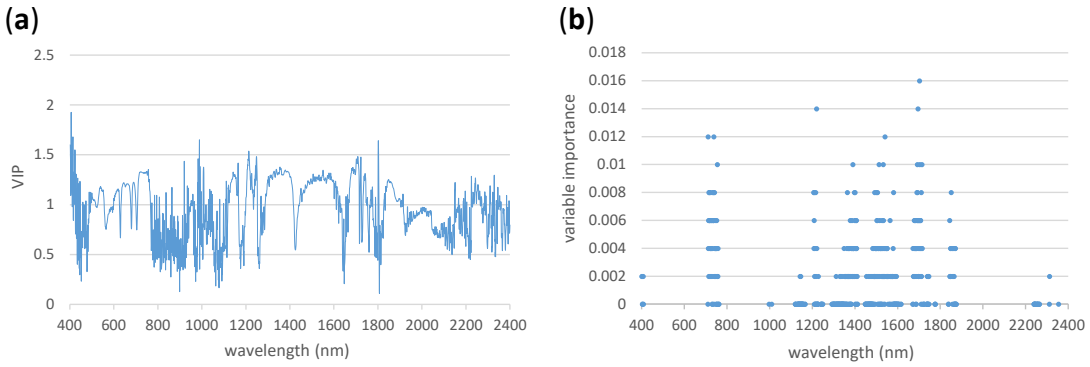



Figure 3. 9 Variable importance for first derivative reflectance computed by PLSR (**a**), and random forest regression (RFR) based on Spearman correlation-selected variables (**b**).

3.3.4.3 Second Derivative

The important 2 D variables computed based on VIP (Figure 3.10a) are even more discrete across the entire spectrum compared to those for 1 D reflectance. This implies that there are relatively few variables relevant to the variation of $\Psi_{stem}$ values in this feature group. As the VIP values of several regions are close to zero, the significant spectral regions can be determined more clearly. Interestingly, the evident regions based on VIP values are quite similar, to those selected by the Spearman correlation. These regions are at around 650–750, 1155, 1370–1420, 1720, and 1870 nm (Figure 3.10b). It seems the regions at around 700 and 1410 were relatively more significant according to the variable importance computed by RFR.

(**a**) (**b**)

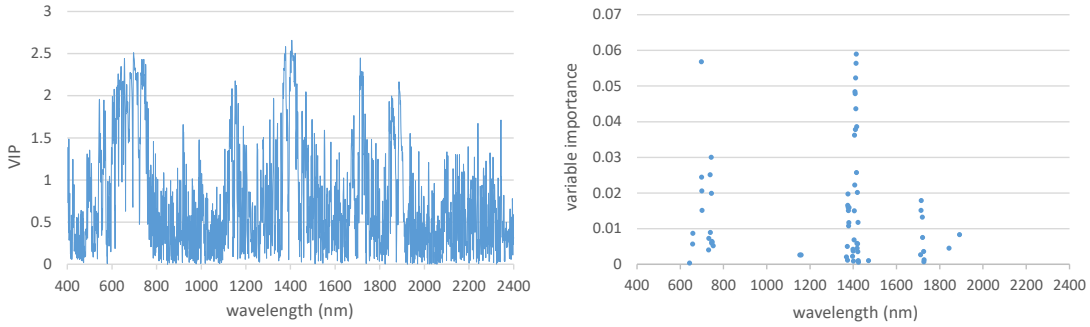



Figure 3. 10 Variable importance for second derivative reflectance computed by PLSR (**a**), and RFR based on Spearman correlation-selected variables (**b**).

3.3.4.4 Continuum Removal Features

With CR variables, both RFR and SVR performed better when using the full set of variables. The top five important variables ranked by PLSR and RFR are the same (Figure 3.11), including continuum slope of the region centered at 670 nm, continuum slope of the region centered at 1925 nm, absorption area of the region centered at 670 nm, full width at half maximum of band depth (FWHM) of the region centered at 670 nm, and continuum slope of the region centered at 1440 nm.

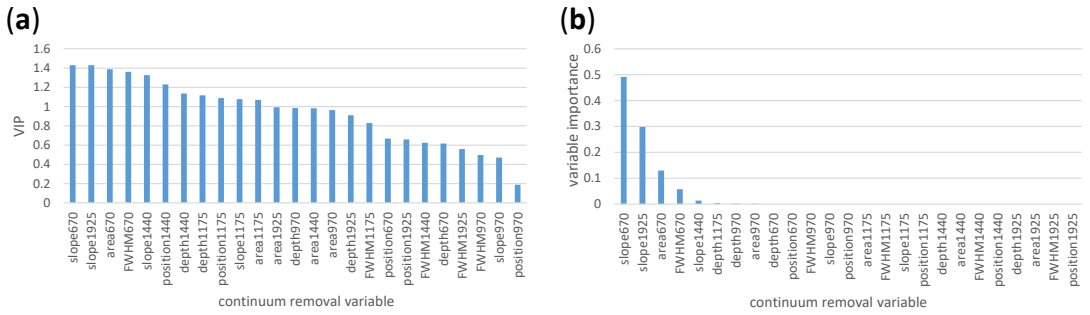



Figure 3. 11 Variable importance for continuum removal variables computed by PLSR (**a**), and RFR (**b**). FWHM refers to width at half maximum of band depth in the target bands. The numbers refer to the center wavelength of the target bands.

3.3.4.5 Simple Ratio Indices

Most of the SI variables make similar contributions to the modeling according to the nondrastically different VIP values in the heatmap (Figure 3.12a). Several important regions can be identified based on the strip-like differences of color in the heatmap. They include the regions of 500–650, 700–730, 1400, 1700, 1900, and 2000–2400 nm. Some indices only play a significant role when formed by the reflectance of adjacent wavelengths, so there are evident differences of color in the heatmap close to the 1:1 line, involving 1050–1150, 1200–1300, and 1800 nm. RFECV-selected variables have similar significance based on absolute coefficient values computed by SVR (Figure 3.12b). Most of the selected indices are built based on the reflectance close to each other by wavelengths, and thus, the colored regions go along the 1:1 line. These regions mainly include 400–500, 750–1300, 1500–1850, 1900–2400 nm. Besides, some strip-like color bars indicate some selected indices are calculated by the ratios of reflectance between 400–600 and 700 nm, 1750–1850 and 1400 nm, as well as 2250–2350 and 1900 nm.

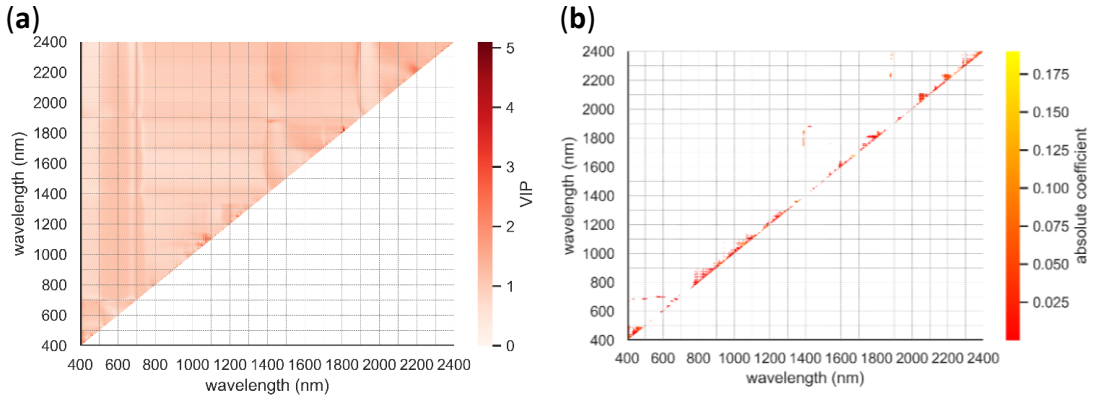



Figure 3. 12 Variable importance for simple ratio indices computed by PLSR (**a**) and SVR based on RFECV-selected variables (**b**).

### 3.4 Discussion

This study attempts to establish a quantitative correlation between the spectral reflectance of vine canopy leaves in the VIS, NIR, and SWIR regions of the spectrum, and the $\Psi_{stem}$, which is used as a proxy for GWS. Since this relationship has been reported to be affected by numerous factors, including growing conditions, phenological stages (Wang & Li, 2012), leaf homogeneity (Seelig et al., 2008), cultivars (Gutiérrez-Gamboa et al., 2019), leaf age (Guyot, 1990), and leaf position (Turner, 1988), the study trial was set up using the same cultivar with similar age and consistent sampling method throughout the field sampling to help minimize these factors. However, in situ heterogeneity of soil type and microclimate conditions will influence growth patterns and/or water stress levels within and between grapevines, so sampling was undertaken frequently over the critical period (late November to early February) to capture sufficient variability for analysis. Note that the number of data collected for this study is small (n = 85). In addition, these are commercial vineyards subjected to normal management practices, and one of the objectives of this study was to ensure that the collected data represented vine responses to these conditions as much as possible.

#### *3.4.1 The Effects of Data Transformation on the Estimation of Grapevine Water Status*

The models fitted using the SI outperformed those regressed with the other transformed data among pipelines with numbers 1–5 and among pipelines with numbers 6–37 in Table 3.4. When the input variables were augmented from 2001 reflectance values over 400–2400 nm to 2,001,000 ratio values, new information was produced, which largely increased the potential of correlation with $\Psi_{stem}$. It is observed 1D, 2D, CR transformations do not improve the modeling performance compared to models trained with raw reflectance data among pipelines. One possible reason for the poor performance of pipelines with CR preprocessing is that since the five selected spectral regions (560–750, 900–1060, 1080–1250, 1280–1660, and 1830–2210 nm) did not fully cover the entire study spectrum of 400–2400 nm, the rejected regions may have contained sufficient 'diffuse' information to affect modeling performance when correlating reflectance with $\Psi_{stem}$ values. For this work, a plant probe with a stable incident angle and light intensity was used in contact with leaves to provide standardized survey conditions as described in Section 3.2.4. It minimized the influences, of illumination conditions, angle of the sun, and background

interference on in situ spectral measurements. That is why original reflectance in this study can achieve high accuracy of $\Psi_{stem}$ estimation, which supports the study of González-Fernández et al. (2019).

### *3.4.2 Significantly Important Spectral Regions Derived from Variable Selection*

Due to multicollinearity within the hyperspectral reflectance, the removal of noncontributing variables is problematic. However, it is important to remove less informative variables to help minimize modeling noise by these variables (Xiaobo et al., 2010). It is observed that Spearman correlation works better using 1D and 2D reflectance as input variables, and RFECV performs better using raw reflectance and SI as input variables. Raw reflectance and SI have much higher multicollinearity than 1D and 2D do according to the VIP values (Figures 3.8a, 3.9a, 3.10a, and 3.12a). Since PLSR extracts the shared variance between predictor variables, the higher VIP values compared to predictor variables, the less correlation is between them. This observation is similar to the findings of Bhadra et al. (2020). However, they used Pearson correlation and RFR instead of Spearman correlation and RFECV to select variables. The full set of CR variables was selected, because this feature group is not highly dimensional, variable elimination may remove some relevant, but diffuse, variables, and thus, reduce modeling performance.

The important regions were those bounded by the spectral bands that were determined by the modeling pipelines with the highest estimation accuracy for each feature group. In the VIS spectrum, the important bands identified are 400–430 nm and 650–750 nm. The 400–430 nm band corresponds to the blue band representing strong absorption by chlorophyll-a, chlorophyll-b, and carotenoids (carotenes and xanthophylls). Variations in these compounds indicate cumulative effects of water stress, and are, thus, indirectly related to variation in GWS, and thus, $\Psi_{stem}$ values (Zarco-Tejada et al., 2013). The 650–750 nm corresponds to red and red edge bands and describes the concentrations (Ballester et al., 2018) and ratios (Zovko et al., 2019) of chlorophyll-a and -b, again indicating water stress status. The ranking of CR features indicates the absorption band at around 670 nm (slope, area, and full width at half maximum of band depth (FWHM)) is significant. These CR variables describe the shape of the absorption curve within this band. This has been observed in previous studies (Blackburn, 2007; Campbell et al., 2007), which found that the position and shape of the absorption curve in the red edge band changed, due to water stress-induced changes of chlorophyll content in the vine leaves.

In the NIR spectrum, the important bands determined are 800–1250 nm. Within this band, there is a partial reflectance response to two weak water absorption bands at 970 and 1200 nm (Woolley, 1971). Variation of reflectance in this band can also be related to changes in internal leaf structure resulting from dehydration (Jacquemoud & Baret, 1990) and the decomposition of celluloses and proteins, due to water stress (Yin et al., 2017).

In the SWIR spectrum, the important bands identified are 1370–1420, 1500–1595, 1700–1720, 1850–1890, 2050–2370 nm. Reflectance responses in the SWIR spectrum are partially determined by dry leaf matter (i.e., lignin, cellulose, and protein), but mainly by two strong water absorption features centered at 1400 and 1940 nm (Curran, 1989). Extra consideration should be given when using water absorption bands (1400 and 1940 nm) to estimate plant water status. The spectral reflectance acquired in this study was a contact measurement using a leaf clip and contact probe with artificial illumination. However, if GWS is estimated by airborne or

space-borne data, the reflectance at 1400 and 1940 nm would become useless because the solar energy, as source illumination, is largely absorbed by atmospheric water vapor before reaching the surface of the earth. As dry leaf matter remains relatively constant under low water deficit regimes, water content is considered the dominant factor influencing the SWIR spectrum from 1300 to 2500 nm. However, as water stress increases and leaf water content declines, the effect of dry leaf matter on spectral reflectance becomes more apparent (Gao & Goetz, 1994). Disregarding the water absorption regions (1400 and 1940 nm), the remaining bands in the SWIR spectrum agree well with the findings of (Sims & Gamon, 2003) (1520–1540 nm) and (Cheng et al., 2011; Thulin et al., 2014; Tian et al., 2001) (1650–1850 and 2000–2270 nm). For the CR feature group, bands (1280–1660, and 1830–2210 nm) were selected based on their continuum slope. This was calculated based on the ratio of the reflectance difference to the bandwidth. As the bandwidth was predetermined, the significance of continuum slope can be attributed to reflectance difference and suggests the potential to use reflectance differences as variables for estimating $\Psi_{stem}$.

The ensemble of selected variables in this study includes RFECV-selected variables from raw reflectance, Spearman correlation-selected variables from 1D reflectance, Spearman correlation-selected variables from 2D reflectance, full set of CR variables, and RFECV-selected variables from SI. This method, although not resulting in evident improvement of estimation accuracy, generated the highest $R^2$ of 0.79 and lowest RMSE of 128 on the test set among pipelines 6–37 in Table 3.4. This proves the benefit offered by ensemble technique beyond what can be achieved by a single combination of a data transformation technique and a variable selection method, which was proposed by Feilhauer et al. (2015).

### *3.4.3 The Performance of Regression Models*

Although previous studies have stated that the NIR-SWIR spectrum is more suitable for water status estimation (Rallo et al., 2014; Rapaport et al., 2015), this paper suggests that statistically significant wavelengths correlated with $\Psi_{stem}$ variation span several spectral regions over the entire spectrum, when different transformed datasets are used as inputs, in the modeling pipelines employed by this study. The poor performance of VI also implies the limitation of using the reflectance given at two to three wavelengths. This was in agreement with the study of Feilhauer et al. (2015), which stated the spectral features related to biochemical indicators were dispersed across multiple bands, and thus, needed to be considered collectively. Therefore, the multivariate techniques that were utilized in this study attempted to make the best use of the entire hyperspectral spectrum instead of focusing on conventional indices derived from reflectance at two or three wavelengths. The advantage of this way has been demonstrated by Romero et al. (2018). PLSR, RFR, and SVR were regressed with either the entire spectrum or a subset of this spectrum. Despite the high dimensionality and multicollinearity inherent in hyperspectral data, PLSR can reduce this complexity down to a few independent variables and attained the best accuracy for $\Psi_{stem}$ estimation. One explanation is that PLSR can effectively integrate the shared variance of both directly and indirectly relevant bands, as explained in Section 4.2. This capability has also been observed by several studies using the same technique to estimate crop water status from hyperspectral data (Cassel et al., 1999; Colombo et al., 2008; Das et al., 2021; Rallo et al., 2014). Since PLSR simulates a linear relationship, this suggests that there is a linear relationship between the extracted information from important bands and $\Psi_{stem}$ values for all the

feature groups except 2D reflectance. One possible explanation is that its relationship with $\Psi_{stem}$ may not be best described by a linear model, such as PLSR. Since RFR and SVR are able, to simulate nonlinearity, this may be the reason why these two models can outperform PLSR when using 2D reflectance as the input dataset. Reduced variable size resulting from variable selection improves the performance of RFR and SVR for most of the feature groups except for CR variables. RFR and SVR were reported to show robustness on high-dimensional data (Doktor et al., 2014; Stamenkovic et al., 2013; Wang et al., 2016). However, this study demonstrates the advantages of variable selection in terms of increasing model performance by RFR and SVR. This may be attributed to the decreased collinearity in the reduced input data. SVR has been found (in other studies (Axelsson et al., 2013)) to suffer from multicollinearity when fitted using the full-spectrum datasets, with improved performance when the most informative bands were used as input data instead.

### 3.5 Conclusions

This study investigated the relationship between stem water potential ($\Psi_{stem}$) and leaf-scale hyperspectral reflectance (400–2400 nm) collected between late November 2020 and early February 2021 from two New Zealand vineyards using 38 modeling pipelines. These pipelines show that partial least squares regression trained with simple ratio indices based on the entire spectrum provided the best $\Psi_{stem}$ predictions ($R^2$ = 0.85; RMSE = 110 kPa), significantly outperforming the linear regression using classical vegetation index as an input variable. Additional results reveal the benefit of increasing accuracy at $\Psi_{stem}$ prediction using an ensemble of selected variables composed of multicombination of transformed data and variable selection methods. The above-mentioned outcomes can be used to tailor an automatic data processing and modeling pipeline to estimate $\Psi_{stem}$. Accordingly, if sufficient hyperspectral measurements are undertaken, this will provide a means of delivering a rapid and nondestructive estimation of $\Psi_{stem}$, and thus, grapevine water status. Information on individual grapevine water status should enable vineyards to tailor irrigation on a per vine basis rather than a per block status, which is the general practice at present. This would enable better control of the desirable traits in the grape berries, that are affected by water content. This would improve wine quality, and therefore, the price achieved. Due to the limited data size obtained in this study, future studies can potentially focus on validating these pipelines using more samples collected from different growing stages and years. Extra consideration should be taken when using airborne or space-borne imaging for estimation, because of water absorption bands. However, more research is required to be carried out, and this study has proved a concept of estimating grapevine water status using a ground-based hyperspectral spectroradiometer.

**References**


Acevedo-Opazo, C., Tisseyre, B., Guillaume, S., & Ojeda, H. (2008). The potential of high spatial resolution information to define within-vineyard zones related to vine water status. *Precision Agriculture, 9*(5), 285-302. https://doi.org/10.1007/s11119-008-9073-1

Axelsson, C., Skidmore, A. K., Schlerf, M., Fauzi, A., & Verhoef, W. (2013). Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. *International Journal of Remote Sensing, 34*(5), 1724-1743.

Ballester, C., Zarco-Tejada, P., Nicolás, E., Alarcón, J. J., Fereres, E., Intrigliolo, D. S., & Gonzalez-Dugo, V. (2018). Evaluating the performance of xanthophyll, chlorophyll and structure-sensitive spectral indices to detect water stress in five fruit tree species. *Precision Agriculture, 19*(1), 178-193.

Bhadra, S., Sagan, V., Maimaitijiang, M., Maimaitiyiming, M., Newcomb, M., Shakoor, N., & Mockler, T. C. (2020). Quantifying leaf chlorophyll concentration of sorghum from hyperspectral data using derivative calculus and machine learning. *Remote Sensing, 12*(13), 2082.

Blackburn, G. A. (2007). Wavelet decomposition of hyperspectral data: a novel approach to quantifying pigment concentrations in vegetation. *International Journal of Remote Sensing, 28*(12), 2831-2855. https://doi.org/10.1080/01431160600928625

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Campbell, P. K., Middleton, E. M., McMurtrey, J. E., Corp, L. A., & Chappelle, E. W. (2007, May-Jun). Assessment of vegetation stress using reflectance or fluorescence measurements. *J Environ Qual, 36*(3), 832-845. https://doi.org/10.2134/jeq2005.0396

Cao, Z., Wang, Q., & Zheng, C. (2015). Best hyperspectral indices for tracing leaf water status as determined from leaf dehydration experiments. *Ecological indicators, 54*, 96-107.

Cassel, C., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of applied statistics, 26*(4), 435-446.

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering, 40*(1), 16-28.

Chaves, M., Zarrouk, O., Francisco, R., Costa, J., Santos, T., Regalado, A., Rodrigues, M., & Lopes, C. (2010). Grapevine under deficit irrigation: hints from physiological and molecular data. *Annals of Botany, 105*(5), 661-676.

Cheng, T., Rivard, B., & Sánchez-Azofeifa, A. (2011). Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sensing of Environment, 115(2), 659-670. https://doi.org/10.1016/j.rse.2010.11.001

Choné, X. (2001). Stem Water Potential is a Sensitive Indicator of Grapevine Water Status. *Annals of Botany, 87*(4), 477-483. https://doi.org/10.1006/anbo.2000.1361

Colombo, R., Meroni, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., & Panigada, C. (2008). Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. *Remote Sensing of Environment, 112*(4), 1820-1834.

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment, 30*(3), 271-278.

Das, B., Sahoo, R. N., Pargal, S., Krishna, G., Verma, R., Viswanathan, C., Sehgal, V. K., & Gupta, V. K. (2021). Evaluation of different water absorption bands, indices and multivariate models for water-deficit stress monitoring in rice using visible-near infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 247*, 119104.

Demetriades-Shah, T. H., Steven, M. D., & Clark, J. A. (1990). High resolution derivative spectra in remote sensing. *Remote Sensing of Environment, 33*(1), 55-64.

Doktor, D., Lausch, A., Spengler, D., & Thurner, M. (2014). Extraction of Plant Physiological Status from Hyperspectral Signatures Using Machine Learning Methods. *Remote Sensing, 6*(12), 12247-12274. https://doi.org/10.3390/rs61212247

Du, P., Xia, J., Chanussot, J., & He, X. (2012). Hyperspectral remote sensing image classification based on the integration of support vector machine and random forest. 2012 IEEE International Geoscience and Remote Sensing Symposium,

Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., & Liu, S. (2012). Multiple Classifier System for Remote Sensing Image Classification: A Review. *Sensors, 12*(4), 4764-4792. https://www.mdpi.com/1424-8220/12/4/4764

Eitel, J. U., Gessler, P. E., Smith, A. M., & Robberecht, R. (2006). Suitability of existing and novel spectral indices to remotely detect water stress in Populus spp. *Forest Ecology and Management, 229*(1-3), 170-182.

Elsayed, S., Mistele, B., & Schmidhalter, U. J. F. P. B. (2011). Can changes in leaf water potential be assessed spectrally? *, 38*(6), 523-533.

Engler, R., Waser, L. T., Zimmermann, N. E., Schaub, M., Berdos, S., Ginzler, C., & Psomas, A. (2013). Combining ensemble modeling and remote sensing for

mapping individual tree species at high spatial resolution. *Forest Ecology and Management, 310*, 64-73.

Eriksson, L., Byrne, T., Johansson, E., Trygg, J., & Vikström, C. (2013). *Multi-and megavariate data analysis basic principles and applications* (Vol. 1). Umetrics Academy.

Feilhauer, H., Asner, G. P., & Martin, R. E. (2015). Multi-method ensemble selection of spectral bands related to leaf biochemistry. *Remote Sensing of Environment, 164*, 57-65.

Gamon, J., Penuelas, J., & Field, C. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment, 41*(1), 35-44.

Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment, 58*(3), 257-266.

Gao, B.-C., & Goetz, A. F. (1994). Extraction of dry leaf spectral features from reflectance spectra of green vegetation. *Remote Sensing of Environment, 47*(3), 369-374.

Gausman, H. W. (1974). Leaf reflectance of near-infrared. *Photogrammetric Engineering, 40*(2), 183-191.

González-Fernández, A. B., Rodríguez-Pérez, J. R., Marabel, M., & Álvarez-Taboada, F. (2015). Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression. *Scientia Horticulturae, 188*, 15-22.

González-Fernández, A. B., Sanz-Ablanedo, E., Gabella, V. M., García-Fernández, M., & Rodríguez-Pérez, J. R. (2019). Field Spectroscopy: A Non-Destructive Technique for Estimating Water Status in Vineyards. *Agronomy, 9*(8), 427.

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and computing, 27*(3), 659-678.

Gutiérrez-Gamboa, G., Pérez-Donoso, A. G., Pou-Mir, A., Acevedo-Opazo, C., & Valdés-Góme, H. (2019). Hydric behaviour and gas exchange in different grapevine varieties (Vitis vinifera L.) from the Maule Valley (Chile). *South African Journal of Enology and Viticulture, 40*(2), 1-1.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research, 3*(Mar), 1157-1182.

Guyot, G. (1990). Optical properties of vegetation canopies. *Optical properties of*

*vegetation canopies.*, 19-43.

Hua, J., Xiong, Z., Lowey, J., Suh, E., & Dougherty, E. R. (2005). Optimal number of features as a function of sample size for various classification rules. *Bioinformatics, 21*(8), 1509-1515.

Hunt, E. R. J., & Rock, B. N. (1989). Detection of changes in leaf water content using near-and middle-infrared reflectances. *Remote Sensing of Environment, 30*(1), 43-54.

Jacquemoud, S., & Baret, F. (1990). PROSPECT: a model of leaf optical properties spectra. *Remote Sensing of Environment, 34*, 75-91.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology, 50*(4), 663-666.

Kokaly, R. F., & Clark, R. N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment, 67*(3), 267-287.

Krishna, G., Sahoo, R. N., Singh, P., Bajpai, V., Patra, H., Kumar, S., Dandapani, R., Gupta, V. K., Viswanathan, C., & Ahmad, T. (2019). Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing. *Agricultural Water Management, 213*, 231-244.

Lavoie-Lamoureux, A., Sacco, D., Risse, P. A., & Lovisolo, C. (2017, Apr). Factors influencing stomatal conductance in response to water availability in grapevine: a meta-analysis. *Physiol Plant, 159*(4), 468-482. https://doi.org/10.1111/ppl.12530

Le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailler, J.-Y., Bréda, N., Genet, H., Davi, H., & Dufrêne, E. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment, 112*(10), 3846-3864.

Li, X., Zhang, Y., Bao, Y., Luo, J., Jin, X., Xu, X., Song, X., & Yang, G. (2014). Exploring the best hyperspectral features for LAI estimation using partial least squares regression. *Remote Sensing, 6*(7), 6221-6241.

Loggenberg, K., Strever, A., Greyling, B., & Poona, N. (2018). Modelling water stress in a shiraz vineyard using hyperspectral imaging and machine learning. *Remote Sensing, 10*(2), 202.

Maimaitiyiming, M., Ghulam, A., Bozzolo, A., Wilkins, J. L., & Kwasniewski, M. T. (2017). Early detection of plant physiological responses to different levels of water stress using reflectance spectroscopy. *Remote Sensing, 9*(7), 745.

Martens, H., & Naes, T. (1992). *Multivariate calibration*. John Wiley & Sons.

Ojeda, H., Andary, C., Kraeva, E., Carbonneau, A., & Deloire, A. (2002). Influence of pre- and postveraison water deficit on synthesis and concentration of skin phenolic compounds during berry growth of Vitis vinifera cv. Shiraz. *American Journal of Enology and Viticulture, 53*(4), 261-267.

Ojeda, H., Carrillo, N., Deis, L., Tisseyre, B., Heywang, M., & Carbonneau, A. (2005, 23-27 August). *Precision viticulture and water status II: Quantitative and qualitative performance of different within field zones, defined from water potential mapping* XIV International GESCO Viticulture Congress, Geisenheim, Germany.

Oumar, Z., & Mutanga, O. (2010). Predicting plant water content in Eucalyptus grandis forest stands in KwaZulu-Natal, South Africa using field spectra resampled to the Sumbandila Satellite Sensor. *International Journal of Applied Earth Observation and Geoinformation, 12*(3), 158-164.

Penuelas, J., Gamon, J. A., Griffin, K. L., & Field, C. B. J. R. S. o. E. (1993). Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. *46*(2), 110-118.

Pôças, I., Gonçalves, J., Costa, P. M., Gonçalves, I., Pereira, L. S., & Cunha, M. (2017). Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region. *International Journal of Applied Earth Observation and Geoinformation, 58*, 177-190.

Pôças, I., Rodrigues, A., Gonçalves, S., Costa, P., Gonçalves, I., Pereira, L., & Cunha, M. (2015). Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices. *Remote Sensing, 7*(12), 16460-16479. https://doi.org/10.3390/rs71215835

Rallo, G., Minacapilli, M., Ciraolo, G., & Provenzano, G. (2014). Detecting crop water status in mature olive groves using vegetation spectral measurements. *Biosystems Engineering, 128*, 52-68. https://doi.org/10.1016/j.biosystemseng.2014.08.012

Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing, 109*, 88-97. https://doi.org/10.1016/j.isprsjprs.2015.09.003

Rienth, M., & Scholasch, T. (2019). State-of-the-art of tools and methods to assess vine water status. *OENO One, 53*(4), 619-637.

Rodríguez-Pérez, J. R., Ordóñez, C., González-Fernández, A. B., Sanz-Ablanedo, E., Valenciano, J. B., & Marcelo, V. (2018). Leaf water content estimation by functional linear regression of field spectroscopy data. *Biosystems Engineering, 165*, 36-46.

Rodríguez-Pérez, J. R., Riaño, D., Carlisle, E., Ustin, S., & Smart, D. R. (2007). Evaluation of hyperspectral reflectance indexes to detect grapevine water status in vineyards. *American Journal of Enology and Viticulture, 58*(3), 302-317.

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. https://doi.org/10.1016/j.compag.2018.02.013

Seelig, H.-D., Adams, W. W., Hoehn, A., Stodieck, L. S., Klaus, D. M., & Emery, W. J. (2008). Extraneous variables and their influence on reflectance-based measurements of leaf water content. *Irrigation Science, 26*(5), 407-414.

Seijo-Pardo, B., Bolón-Canedo, V., & Alonso-Betanzos, A. (2019). On developing an automatic threshold applied to feature selection ensembles. *Information Fusion, 45*, 227-245.

Sims, D. A., & Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment, 84*(4), 526-537.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing, 14*(3), 199-222.

Stamenkovic, J., Tuia, D., De Morsier, F., Borgeaud, M., & Thiran, J.-P. (2013). Estimation of soil moisture from airborne hyperspectral imagery with support vector regression. 2013 5th workshop on hyperspectral image and signal processing: evolution in remote sensing (WHISPERS),

Strachan, I. B., Pattey, E., & Boisvert, J. B. (2002). Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. *Remote Sensing of Environment, 80*(2), 213-224.

Thulin, S., Hill, M. J., Held, A., Jones, S., & Woodgate, P. (2014). Predicting levels of crude protein, digestibility, lignin and cellulose in temperate pastures using hyperspectral image data. American Journal of Plant Sciences, 5, 997-1019.

Tian, Q.-j., Tong, Q.-x., Pu, R.-l., Guo, X., & Zhao, C. (2001). Spectroscopic determination

of wheat water status using 1650-1850 nm spectral absorption features. International Journal of Remote Sensing, 22(12), 2329-2338.

Tosin, R., Pocas, I., Gonçalves, I., & Cunha, M. (2020). Estimation of grapevine predawn leaf water potential based on hyperspectral reflectance data in Douro wine region. *Vitis, 59*, 9-18.

Turner, N. C. (1988). Measurement of plant water status by the pressure chamber technique. *Irrigation Science, 9*(4), 289-308.

Van Leeuwen, C., Trégoat, O., Choné, X., Bois, B., Pernet, D., & Gaudillère, J.-P. (2009). Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? *OENO One, 43*(3), 121-134.

Vasques, G., Grunwald, S., & Sickman, J. (2008). Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. *Geoderma, 146*(1-2), 14-25.

Wang, L. a., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal, 4*(3), 212-219.

Wang, Q., & Li, P. (2012). Identification of robust hyperspectral indices on forest leaf water content using PROSPECT simulated dataset and field reflectance measurements. *Hydrological Processes, 26*(8), 1230-1241.

Woolley, J. T. (1971). Reflectance and transmittance of light by leaves. *Plant Physiology, 47*(5), 656-662.

Xiaobo, Z., Jiewen, Z., Povey, M. J., Holmes, M., & Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. *Analytica chimica acta, 667*(1-2), 14-32.

Xu, M., Liu, H., Beck, R., Lekki, J., Yang, B., Shu, S., Kang, E. L., Anderson, R., Johansen, R., & Emery, E. (2019). A spectral space partition guided ensemble method for retrieving chlorophyll-a concentration in inland waters from Sentinel-2A satellite imagery. *Journal of Great Lakes Research, 45*(3), 454-465.

Yin, W., Zhang, C., Zhu, H., Zhao, Y., & He, Y. (2017). Application of near-infrared hyperspectral imaging to discriminate different geographical origins of Chinese wolfberries. *PLoS One, 12*(7), e0180534. https://doi.org/10.1371/journal.pone.0180534

Zarco-Tejada, P. J., González-Dugo, V., Williams, L. E., Suárez, L., Berni, J. A. J., Goldhamer, D., & Fereres, E. (2013). A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band

airborne imagery and the CWSI thermal index. *Remote Sensing of Environment, 138*, 38-50. https://doi.org/10.1016/j.rse.2013.07.024

Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., & Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 39*(7), 1491-1507.

Zarco-Tejada, P. J., Pushnik, J., Dobrowski, S., & Ustin, S. (2003). Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sensing of Environment, 84*(2), 283-294.

Zovko, M., Žibrat, U., Knapič, M., Kovačić, M. B., & Romić, D. (2019). Hyperspectral remote sensing of grapevine drought stress. *Precision Agriculture, 20*(2), 335-347.

Chapter 3 investigates important spectral bands that are correlated to GWS variation. The significant bands, although dispersed across the entire spectrum according to the transformation approaches, can be identified that they concentrate around the blue, red, and red edge, two weak water absorption bands in NIR, and two strong water absorption bands and dry matter-relevant bands in SWIR. This result enables the use of these relationships for remotely tracking the hydration status of grapevines from airborne platforms. However, atmospheric vapor limits the usefulness of water absorption bands. Therefore, the next chapter is aimed at exploring the potential environmental variables for improving the monitoring capacities of multispectral UAVs in estimating GWS.

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: | Hsiang-En Wei |
| Name and title of main supervisor: | Dr Miles Grafton, Senior Lecturer |
| In which chapter is the manuscript/published work? | Chapter 4 |
| What percentage of the manuscript/published work was contributed by the student? | 82.5% |

Describe the contribution that the student has made to the manuscript/published work:

The candidate used ancillary variables (vegetation characteristics, temporal trends, weather conditions, and soil/terrain data) to improve the monitoring capabilities of multispectral UAV in estimating grapevine water status and provided insights into the contribution of each ancillary variable.

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:
Wei, H.-E.; Grafton, M.; Bretherton, M.; Irwin, M.; Sandoval, E. 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 Sens. 2022, 14, 5918. https://doi.org/10.3390/rs14235918

( ) **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: | Hsiang-En Wei 數位簽署者：Hsiang-En Wei 日期：2023.04.11 16:46:34 +12'00' | 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: 2023.05.04 21:17:20 +12'00' |

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

# Chapter 4
# 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

**Publication**: Wei, H.-E.; Grafton, M.; Bretherton, M.; Irwin, M.; Sandoval, E. 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 Sens*. **2022**, *14*, 5918. https://doi.org/10.3390/rs14235918

After determining the $\Psi_{stem}$ -relevant spectral regions over the spectrum, the next focus is to address the issue that the lack of SWIR in most UAV-based multispectral sensors. This affects the ability to effectively monitor vegetation water status. Alternatively, the correlation of $\Psi_{stem}$ with environmental variables, including soil/terrain, weather, and temporal information, are measurable variables that can be taken advantage of to assist in describing $\Psi_{stem}$ variation. This chapter is dedicated to evaluating the second hypothesis – The technical shortcomings of multispectral UAV in estimating GWS can be improved by the integration of spatial information collected from different sources (UAV, aircraft, EM38, weather station), with an objective of selecting the set of useful environmental variables while obtaining the contribution they make toward GWS variability.

**Abstract**: Monitoring and management of grapevine water status (GWS) over the critical period between flowering and veraison plays a significant role in producing grapes of premium quality. Although unmanned aerial vehicles (UAVs) can provide efficient mapping across the entire vineyard, most commercial UAV-based multispectral sensors do not contain a shortwave infrared band, which makes the monitoring of GWS problematic. The goal of this study is to explore whether and which of the environmental variables (temporal trends, weather conditions, and soil/terrain data) may improve the accuracy of GWS estimation using multispectral UAV and provide insights into the contribution, in terms of direction and intensity, for each variable contributing to GWS variation. UAV-derived vegetation indices, slope, elevation, apparent electrical conductivity ($EC_a$), weekly or daily weather parameters, and day of the year (DOY) were tested and regressed against stem water potential ($\Psi_{stem}$), measured by a pressure bomb, and used as a proxy for GWS using three ML algorithms (elastic net, random forest regression, and support vector regression). Shapley Additive exPlanations (SHAP) analysis was used to assess the relationship between selected variables and $\Psi_{stem}$. The results indicate that the root mean square error (RMSE) of the transformed chlorophyll absorption reflectance index-based model improved from 213 to 146 kPa when DOY and elevation were included as ancillary inputs. RMSE of the excess green index-based model improved from 221 to 138 kPa when DOY, elevation, $EC_a$, and daily average windspeed were included as ancillary inputs. The support vector regression best described the relationship between $\Psi_{stem}$ and selected predictors. This study has provided proof of the concept for developing GWS estimation models that potentially enhance the monitoring capacities of UAVs for GWS, as well as providing individual GWS mapping at the vineyard scale. This may enable growers to improve irrigation management, leading to controlled vegetative growth and optimized berry quality.

## 4.1 Introduction

Studies have demonstrated that grapevine water status (GWS) is a key factor in berry composition, affecting both vegetative growth and fruit metabolism (Martínez-Lüscher et al., 2014; Ojeda et al., 2002; Van Leeuwen et al., 2009). The berry composition determines the quality at harvest. During the important phenological stages, bloom, and veraison (Acevedo-Opazo et al., 2013), GWS is ideally under controlled water deficit to benefit berry development. This management practice suppresses competition for photosynthetic resources from vegetative growth (Intrigliolo & Castel, 2009). In addition, this husbandry promotes the accumulation of sugar and anthocyanins (Etchebarne et al., 2010) and prevents oxidative damage to canopies resulting from severe water stress (Min et al., 2019). The fluctuating hydration status in vines varies across vineyards even under a homogeneous irrigation scheme (Brillante et al., 2017), and this subsequently leads to variability in vine growth and berry development (Baciocco et al., 2014; Bramley et al., 2011; López-García et al., 2021). To minimize the variation in grape quality across vineyards, it is important to monitor the temporal and spatial variability of GWS while carrying out corresponding practices to keep water status within an optimal range.

Measuring electromagnetic reflectance from plants has become popular because it is associated with physiological status, phenological stage, and other intrinsic variables of the plants (Liu et al., 2016). In combination with the use of an unmanned aerial vehicle (UAV), it can provide high spatial resolution images that help further

derive vegetation indices (VI) for the whole vineyard in an efficient manner. However, the shortwave infrared band (SWIR, 1–2.5 μm), which has been reported to have an important relationship with foliar water content (Bowyer & Danson, 2004), is not included in most of the commercial and low-cost (lower than USD 5000) UAV-based multispectral sensors (Jang et al., 2020). Indices such as water balance index-2 using bands at 0.538 and 1.5 μm are able to identify water stress in grapevines with promising performance (coefficient of determination ($R^2$) of 0.89) (Rapaport et al., 2015). The inaccessibility of SWIR data has confined the VIs to visible and near-infrared (VNIR) bands, which provide a less satisfying estimation of plant water content (Brook et al., 2020). Arevalo-Ramirez et al. (2020) found that the SWIR band can be reconstructed from the reflectance of the VNIR band using two ML models. Jenal et al. (2019) tackled the issue from a hardware perspective by developing a multispectral imaging system for UAVs that covers 0.4–1.7 μm of the spectrum. Kandylakis et al. (2020) collected spectral data from integrated UAV-based multispectral/SWIR sensors ranging between 0.53–1.7 μm and designed a processing pipeline to execute data analysis for estimating GWS.

Another way to compensate for the missing SWIR band is to utilize the integrated response of plant water status to soil moisture availability, atmospheric demand, plant characteristics, and cultivation practices (Choné, 2001). A study supports the view that cultivar is the main driver of GWS under well-watered conditions, while vegetative expression and soil type become more dominant as water deficit increases (Taylor et al., 2010). Weather variables are other important driving factors characterizing the variation of plant water status (Irmak & Mutiibwa, 2010). These drivers induce hydraulic and metabolic signals that trigger water consumption by crops, along with a range of physiological responses to water deficit, consequently shaping the spatial and temporal variation of GWS in vineyards. Grapevine hydration status is often substituted by predawn leaf water potential ($\Psi_{pd}$), midday stem water potential ($\Psi_{stem}$), midday leaf water potential ($\Psi_{leaf}$), and stomatal conductance. Several studies have been implemented to simulate and explore the relationship between measured GWS and ancillary information composed of vegetative, soil/terrain, temporal, and weather variables (Acevedo-Opazo et al., 2008; Acevedo-Opazo et al., 2010; Brillante et al., 2016; Suter et al., 2019; Tang et al., 2022; Taylor et al., 2012).

To model the relationship between GWS and ancillary information, ML techniques offer an attractive alternative due to their ability to model both linear and nonlinear systems (Kuhn & Johnson, 2013). These tools are capable of capturing the most informative relationships between inputs and outputs, providing predictions based on a set of data variables having a similar distribution to the training set. To further understand the relationship between inputs and outputs, the prediction capabilities as well as the contribution each input makes to the prediction are of interest. That is, it is important to identify the correlation, as well as the causation, between predictors and response variables. Manipulation may occur when trained models possess logical causation. Preprocessing can be carried out on the input dataset to reduce repetitive information and dependence between predictor variables. In this way, a reasonable causation can be captured while the prediction capacities of the ML algorithm may be improved. Interpretability tools, such as Shapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017), are able to provide directionality of the relationships and uncover synergistic effects between multiple variables (Lundberg et al., 2020). When paired with SHAP, ML models become more capable as SHAP offers deeper insights into the trained models whose results are

usually limited to incomplete interpretation, such as feature importance or variable weight (Mangalathu et al., 2020).

As vineyard management is strongly linked to the hydration state of vines, an understanding of GWS dynamics and its dominant drivers could be indispensable for viticulture management. The primary goal of this study is to investigate whether the incorporation of environmental information (weather, temporal, and soil/terrain variables) can enhance the GWS monitoring capacity of UAV-based multispectral sensors without a SWIR band in the production context. The secondary goal is to provide insights into the relationship between GWS and weather, temporal, and soil/terrain variables over the flowering-to-veraison period. To the authors' knowledge, this is the first study where GWS monitoring models have been developed in a temperate climate zone in the Southern Hemisphere using environmental predictors from weather, temporal, and soil/terrain aspects and using SHAP for relationship interpretation. The procedures undertaken were (i) exploring which vegetation indices derived from multispectral imagery are strongly correlated with variation of $\Psi_{stem}$, with those identified serving as the core input in later modeling, as well as (ii) modeling changes in $\Psi_{stem}$ based on VIs, temporal, soil/terrain, and weather variables. The ML model with the best predictive performance and logical causal relationships will then be presented and assessed, (iii) understanding which combination of environmental variables should be selected in the best performing model, with their contribution and relationship with $\Psi_{stem}$ evaluated using SHAP analysis.

## 4.2 Materials and Methods

### *4.2.1 The Context of the Study Vineyards and Study Periods*

The study's vineyards are located at Martinborough in the Greater Wellington Region in New Zealand (NZ) (Figure 4.1). The study sites comprise two commercial vineyards owned by Palliser Estate and are named Wharekauhau and Pencarrow. Our study areas within these two vineyards are 6.6 and 6.7 ha, respectively. A S-map online (https://smap.landcareresearch.co.nz/, accessed on 1st April 2022), developed by Landcare Research, was used to provide a basic soil summary (mapped at about 1:50,000) to support the study. There are mainly two types (Glas_8a.2: 60% and Barr_22a.1: 40%) and three types (Barr_22a.1: 50%, Glas_8a.2: 40%, and Waka_26a.1: 10%) of soils across Wharekauhau and Pencarrow, respectively. Glas_8a.2 and Barr_22a.1 have dominant silt texture in their topsoil and subsoil, with gravelly layers from less than 0.45 m to more than 1 m. They are both described as well-drained, with Glas_8a.2 having moderate soil water holding capacity and Barr_22a having high capacity. Waka_26a.1 has silt and clay texture in its topsoil and subsoil, with a gravel content of less than 3% and plant rooting depth extending beyond 1 m. It is described as imperfectly drained and highly vulnerable to water logging, having moderate soil water holding capacity.

Pinot Noir was chosen as the target cultivar in this study due to its requirement for relatively precise irrigation management. The Pinot Noir vines in both vineyards were planted in 1998–2000 and trained with two-cane vertical shoot positioning. Inter- and intra-row planting space is 2.2 × 1.7 m for Wharekauhau and 2.2 × 1.8 m for Pencarrow. The annual growth cycle of grapevine in NZ comprises budburst, shoot growth, flowering (September–November), fruit set, and veraison (December–February), followed by berry development and harvesting (March–May). Cultivation practices, such as shoot thinning, bud rubbing, and leaf plucking, are regularly

conducted from October to December during the growing season. Between flowering and veraison (termed as the critical period in the following sections), the management of GWS is the most critical determinant in final berry quality.

The trials undertaken in this study took place over two growing seasons. The measurement dates avoided rainy days and matched the most critical period for GWS management. The study periods are between 27 November 2020 and 1 February 2021 for the first growing season and between 29 November 2021 and 21 January 2022 for the second growing season (Figure 4.2).

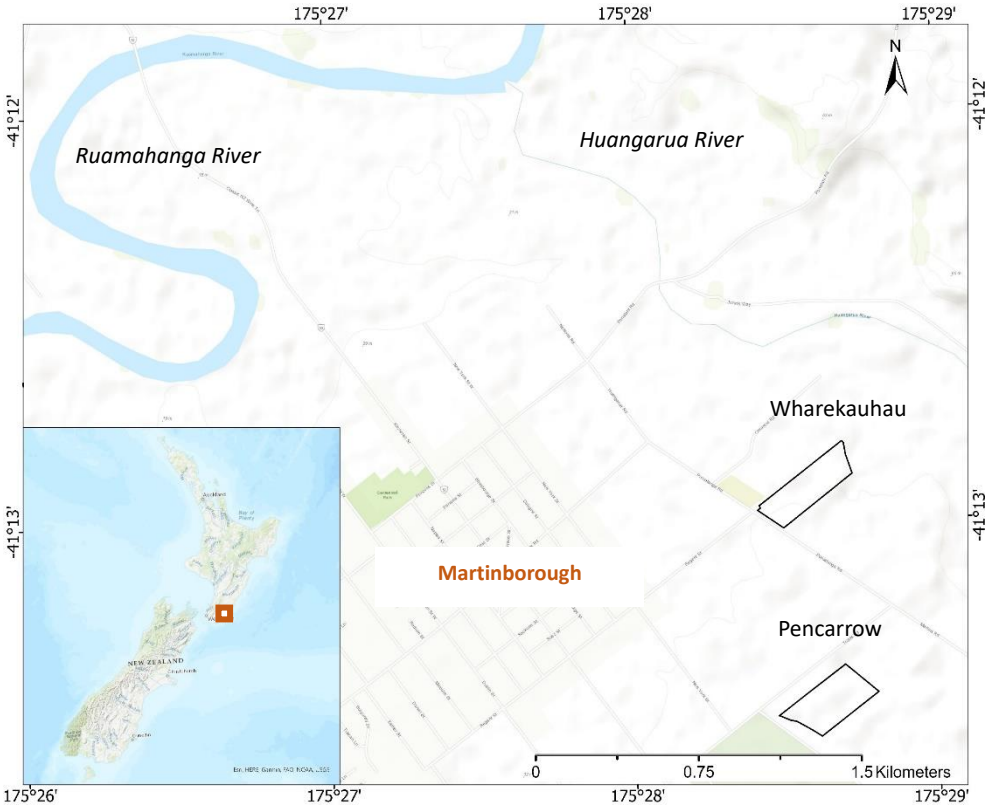



Figure 4. 1 Location of study vineyards.

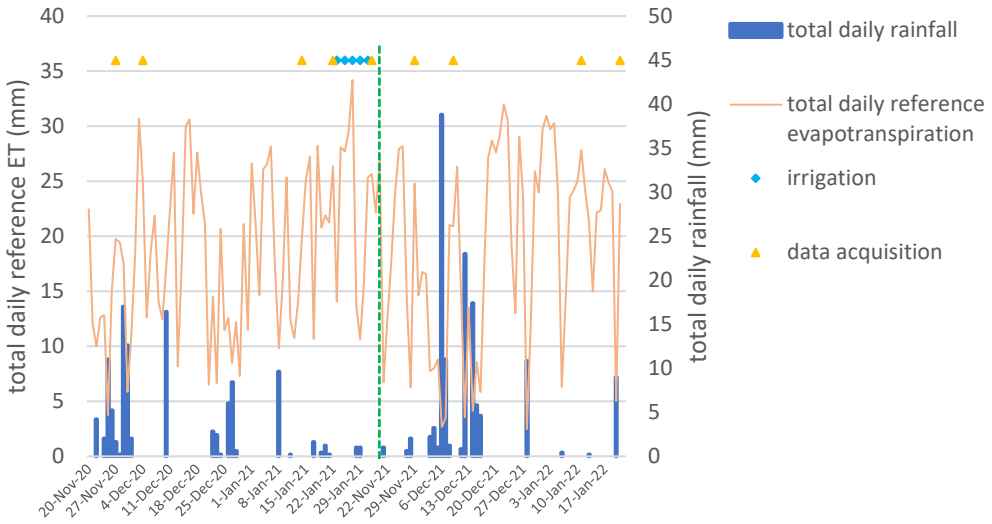


Figure 4. 2 Total daily rainfall recorded by an on-site weather station. Reference evapotranspiration was computed by HARVEST.com (http://harvest.com/, accessed on 1st April 2022) based on the recordings obtained from the weather station. The dates for UAV images and $\Psi_{stem}$ data acquisition were 27 November 2020, 4 December 2020, 14 January 2021, 22 January 2021, 1 February 2021, 29 November 2021, 9 December 2021, 11 January 2022, and 21 January 2022. Irrigation was applied on 23, 25, 27, and 29 January 2021 in the first growing season. The green dash line separates the first and the second growing season.

### *4.2.2 Response Variable-Stem Water Potential*

Stem water potential ($\Psi_{stem}$) was chosen as a proxy for GWS and has been expressed as a comprehensive indicator for early water deficit in vines during the day (Patakas et al., 2005). On each measurement date, several healthy vines were sampled in grids to assess variability across each vineyard (Figure 4.3) using two mature and fully expanded leaves from the middle part of each sampled canopy. The mature and fully expanded leaves are more representative of the status of canopies. A pressure chamber (model: 610, MPS, Albany, NY, USA) was employed between the hours of 13:00 and 15:30 to assess $\Psi_{stem}$ (kPa). Prior to measurement, the sampled leaves were covered with sealable plastic bags for around 1 h. The higher the $\Psi_{stem}$ reading, the more dehydrated the vine. These two measurements (per sampled vine) were averaged to represent the vine's canopy water status. A total of 85 and 63 separate canopies were surveyed in the first and the second growing season (Appendix 1), respectively, and each of their trunk locations was recorded using a global navigation satellite system (GNSS) with real-time kinematic (RTK) correction (model: GPS1200+, Leica Geosystems AG., Heerbrugg, Switzerland).

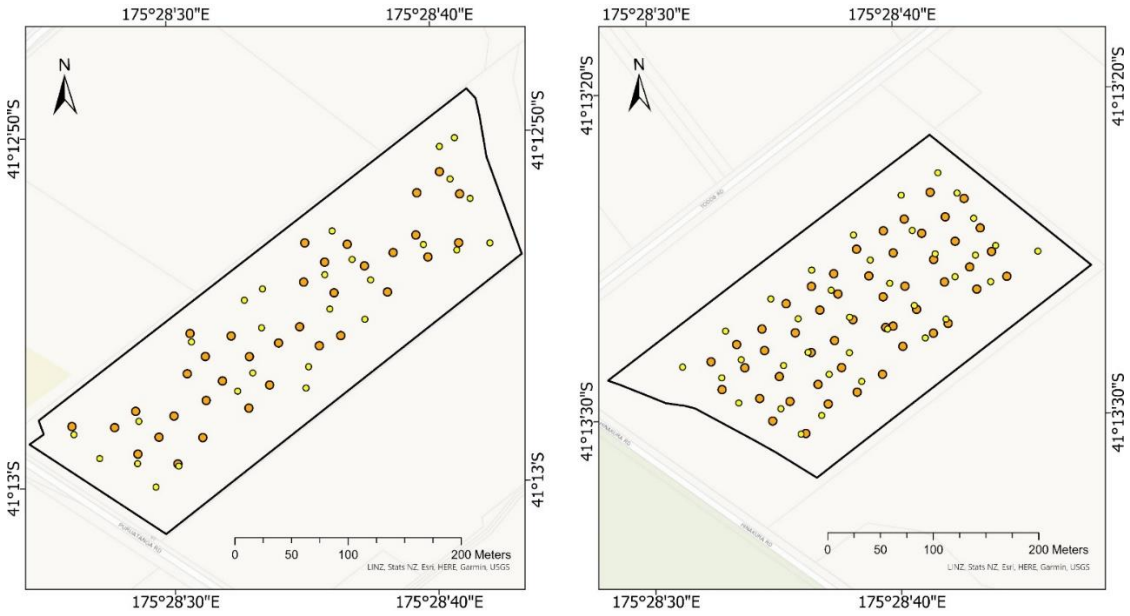



Figure 4. 3 The sampling locations over two study periods. **Left** is the Wharekauhau, and **right** is the Pencarrow vineyards. Orange points represent the observations acquired in the 2020–2021 season, and yellow points represent the observations acquired in the 2021–2022 season.

### *4.2.3 Predictor Variable-Spectral Parameters*

Aerial images were obtained between 11:00 and 13:00 under sunny conditions to minimize the influence of sun angle and shadow and to ensure comparability on the same date that $\Psi_{stem}$ data were measured (Appendix 8). The reflectance, with a spatial resolution of 0.043 m, was recorded by DJI Phantom 4 multispectral UAV (DJI, Shenzhen, China) with six built-in sensors in the blue (450 ± 16 nm), green (560 ± 16 nm), red (650 ± 16 nm), red edge (730 ± 16 nm), and near-infrared (840 ± 26 nm) regions. The DJI Phantom 4 has an integrated sunlight sensor which records irradiance during the flight in the same bands captured by the multispectral sensor for reflectance computing. With this information, UAV images can be normalized, thus allowing for comparison between images taken under different illumination conditions. Photogrammetric processing was applied to the aerial data using Pix4Dmapper (Pix4D SA, Lausanne, Switzerland) to generate digital surface models (DSM), digital terrain models (DTM), and reflectance maps. Settings were designed as follows: key point image was set as full, point cloud densification was set as ½ (optimal), and three for image scale, point density, and minimum number of matches, respectively. Classify point cloud was ticked. Noise filtering and surface smoothing using sharp type were applied. DSMs were produced using an inverse distance weighting algorithm. To increase image spatial accuracy, several ground control points were recorded by GNSS-RTK in each vineyard, and image alignment was subsequently performed in ArcGIS Pro 2.9 (ESRI, Redlands, CA, USA).

The vineyards feature discontinuous vegetation surfaces, so it was necessary to separate canopy pixels from grass and soil pixels to obtain pure information about grapevines. As there is a height difference between grapevine canopies and the surrounding landscape, canopy height was acquired by subtracting DTM from DSM, then creating a binary image with a threshold of 0.9 m to exclude background pixels. For each grapevine, the canopy cordon was set at about 0.8 m from the trunk. Only the vegetation component within 0.5 m distance from the trunk was considered for computing spectral variables in this study since the shoots of adjacent grapevines are often overlapping and intertwined. The acquisition of specific canopy pixels was carried out by overlapping the buffer zones (using recorded trunk location as the center of a circle with a radius of 0.5 m) with the binary raster of canopy height (Figure 4.4). Subsequently, 18 vegetation indices for each sampled grapevine, chosen according to frequency of usage in viticulture (Giovos et al., 2021), and shown in Table 4.1, were calculated based on the mean values of pure canopy pixels using “zonal statistic as table” in ArcGIS Pro.

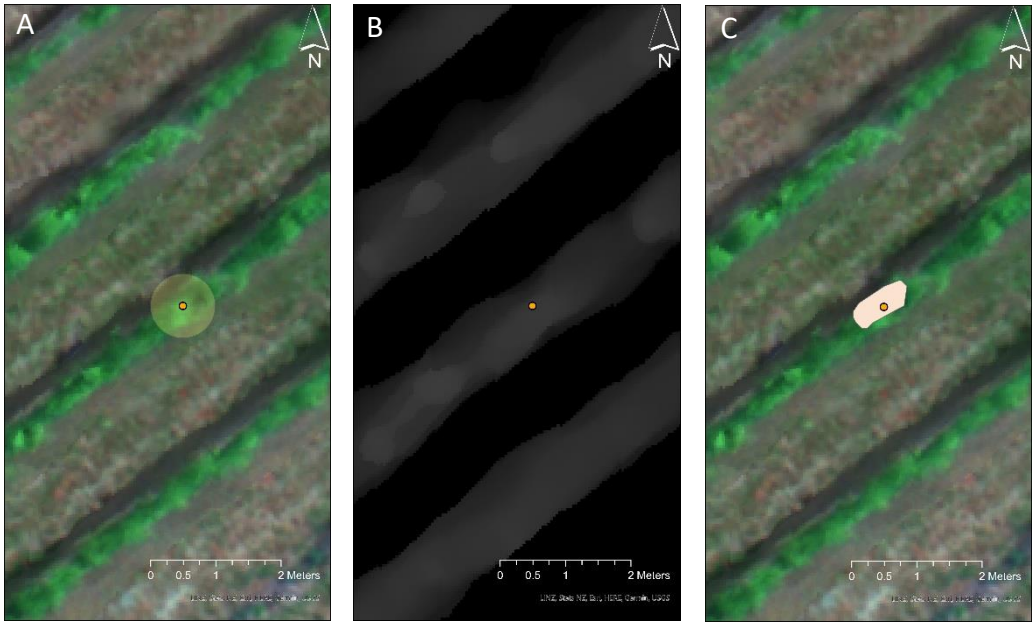



Figure 4. 4 The procedure of getting pure pixels for sampled canopies. (**A**): The subset of an aerial image comprising grapevines and floor vegetation. The buffer zone, the yellow circle, was created according to the recorded location (the orange point) of the sampled grapevine. (**B**): Binary raster of canopy height was generated by subtracting the digital terrain model from the digital surface model. (**C**): Pure grapevine pixels (pale pink region) were acquired by overlapping the buffer zone with the binary raster.

Table 4. 1 List of vegetation indices used in this study.

| Vegetation Index | Acronym | Formula | References |
|---|---|---|---|
| Transformed Chlorophyll Absorption Reflectance Index | TCARI | 3 × ((Red edge−Red) −0.2 × (Red edge−Green) × (Red edge/ Red)) | (Haboudane et al., 2002) |
| Excess Green Index | ExG | 2 × Green−Red−Blue | (Woebbecke et al., 1995) |
| Ratio between Transformed Chlorophyll Absorption Reflectance Index and Optimized Soil Adjusted Vegetation Index | TCARI/OSAVI | - | (Haboudane et al., 2002) |
| Normalized Difference Red Edge Index | NDRE | (NIR−Red edge)/ (NIR + Red edge) | (Barnes et al., 2000) |
| Green Normalized Difference Vegetation Index | GNDVI | (NIR−Green)/ (NIR + Green) | (Gitelson & Merzlyak, 1998) |
| Red Edge Chlorophyll Index | CL red edge | (NIR/ Red edge) −1 | (Gitelson et al., 2005) |
| Modified Triangular Vegetation Index | MTVI1 | 1.2 × (1.2 × (NIR−Green) −2.5 × (Red−Green)) | (Haboudane et al., 2004) |

| | | | |
|---|---|---|---|
| Enhanced Vegetation Index | EVI | $2.5 \times (NIR-Red)/ (NIR + 6 \times Red-7.5 \times Blue + 1)$ | (Huete et al., 2002) |
| Difference Vegetation Index | DVI | NIR−Red | (Tucker, 1979) |
| Modified Soil Adjusted Vegetation Index | MSAVI | $(2 \times NIR + 1- ((2 \times NIR + 1)^2-8 \times (NIR-Red))^{1/2})/ 2$ | (Qi et al., 1994) |
| Simple Ratio | SR | NIR/ Red | (Birth & McVey, 1968) |
| Normalized Difference Vegetation Index | NDVI | (NIR−Red)/ (NIR + Red) | (Rouse et al., 1974) |
| Optimized Soil Adjusted Vegetation Index | OSAVI | (NIR−RED)/ (NIR + Red + 0.16) | (Rondeaux et al., 1996) |
| Normalized Difference Green/Red Index | NGRDI | (Green−Red)/ (Green + Red) | (Tucker, 1979) |
| Red:Green Ratio | R/G index | Red/ Green | (Gamon & Surfus, 1999) |
| Visible Atmospherically Resistant Index | VARI | (Green−Red)/ (Green + Red−Blue) | (Gitelson et al., 2002) |
| Modified Chlorophyll Absorption Ratio Index | MCARI | ((Red edge−Red) −0.2 * (Red edge−Green)) × (Red edge/ Red) | (Daughtry et al., 2000) |
| Canopy volume | - | - | (Ballesteros et al., 2015) |

*4.2.4 Predictor Variable-Soil and Terrain Information*

An EM38-MK2 is an electromagnetic induction (EMI)-based sensor (Geonics Ltd., Mississauga, ON, Canada). The return reading (apparent electrical conductivity ($EC_a$)) is considered a function of soil solid types, soil solution, and soil water content (Cook & Williams, 1998). Spatial patterns of $EC_a$ values have been found to be relatively temporally stable between measurement dates (Heil & Schmidhalter, 2017). Recommended practice is to undertake measurements when soils are near field capacity and $EC_a$-based soil variability is at a maximum (Brevik et al., 2006). In this study, the EM38-MK2 was operated in the vertical dipole mode, with the instrument taking integrated $EC_a$ measurements at about 1.5 m depth. An EMI survey was undertaken on 27 May 2021 by towing the EM38-MK2 at the back of an all-terrain vehicle (less than 0.2 m between EM38-MK2 and the all-terrain vehicle) with a Trimble Yuma tablet incorporating an onboard GPS receiver (model: Yuma, Trimble), accurate to 2–4 m, to geo-reference all point data from the $EC_a$ (mS/m) survey. The total rainfall one week before the survey is 7 mm, and two weeks before the survey is 15.8 mm. The vineyards' infrastructure was confirmed with the grower to ensure there was no interference from buried metal components and perched water tables. $EC_a$ points were measured approximately every 3–10 m along transects and 10 m apart, and values less than 0 mS/m were removed before interpolation. The geostatistical interpolation method, empirical Bayesian kriging (EBK), was used to transform point data onto a continuous surface raster with 1 m resolution in ArcGIS Pro (Appendix 3).

Elevation (m) and slope (degree) information of the location of each sampled canopy was obtained from the 'Wellington LiDAR 1m DEM (2013–2014)' layer provided by the Land Information New Zealand data service (https://data.linz.govt.nz/, accessed on 1st April 2022). This digital elevation model (1 m resolution) was generated by aerial LiDAR captured between 2013 and 2014 for the Greater Wellington region (Appendix 2). For each grapevine, the mean values of $EC_a$, elevation, and slope within 0.5 m distance of the trunk were computed using "zonal statistic as table" in ArcGIS Pro.

*4.2.5 Predictor Variable-Weather and Temporal Data*

Weather data were recorded by an on-site weather station (175.4741, −41.2247 WGS84) established by HARVEST.com (http://harvest.com/, accessed on 1st April 2022). The target variables include air temperature (°C), relative humidity (%), rainfall (mm), wind speed (km/h), and irradiance ($W/m^2$) (Appendix 7). These variables were used to compute mean temperature, mean relative humidity, total rainfall, mean wind speed, and total irradiance based on weekly and daily intervals before each measurement date. It was assumed that climatic conditions were homogeneous across the two vineyards. The temporal data through the season was represented by day of the year (DOY).

*4.2.6 Hierarchical Clustering*

One of the goals of this study was to identify causal relationships that may enable growers to utilize or modify GWS according to tailored quality standards. Accordingly, the models developed in this study were required to not only have predictive capabilities, but also logical causality. Initially, all predictor variables were used for regression against $\Psi_{stem}$ without clustering preprocessing, so that most of the causal relationships were not intuitive. Subsequent hierarchical clustering was employed to generate groups of highly correlated variables and to produce input datasets with less repetitive information. As a result, the model still exhibits similar prediction results using fewer predictors, and all causal relationships are reasonable and intuitive.

The hierarchical clustering for variables was carried out as an exploratory data analysis (Morgenthaler, 2009). Hierarchical clustering was chosen in this study because it can provide a clear overview of the input datasets in terms of grouping while incorporating expert domain knowledge without the need to pre-specify cluster numbers. This is an important step because repetitive information among variables decreases the precision of ML modeling and complicates the interpretation of results. In this study, the similarity between clusters was based on correlation distance, which means that variables which are more positively correlated, based on Pearson correlation, will first merge as one cluster, until all variables are combined into one (James et al., 2013). The criterion used to determine the distance between two clusters, also called linkage, was set to "complete", referring to the largest dissimilarity between observations in clusters. Hierarchical clustering, regression modeling, and SHAP analysis were implemented with Python 3.9.

*4.2.7 Regression Modeling*

On the basis of the clusters generated by hierarchical clustering, combinations of predictor variables were formed by picking one variable from each cluster or not picking any variable from any cluster. All possible combinations, 576 in total, were

investigated along with three types of ML models for performance comparison. Instead of using conventional variable selection methods such as filtering (e.g., Pearson correlation) or wrapper (e.g., recursive feature elimination), all combinations were evaluated, since the behavior of one variable is influenced by the presence of others. Thus, significant variables would not be removed due to weak correlation with the response variable or having less importance in some of the input datasets. The total samples (*n* = 148), collected from both growing seasons, were split into training (*n* = 103) and test (*n* = 45) sets using a 70/30 ratio. This split was carried out and stratified according to the date of measurement to ensure that both training and test sets had corresponding percentages of samples for each date. All predictor variables were standardized to have mean values equivalent to 0 and a standard deviation of 1.

Elastic net (EN), random forest regression (RFR), and support vector regression (SVR) were applied to estimate $\Psi_{stem}$ based on weather, soil/terrain, temporal, and spectral variables. As the performance of regression models is influenced by their hyperparameters, it was necessary to tune the hyperparameters beforehand to prevent overfitting. This enabled the regression algorithms to exploit their potential. Grid searching on the training set with 10-fold cross-validation was used to search for the best combination of hyperparameters. A list of tuned hyperparameters and their ranges for each algorithm is displayed in Table 4.2. The test dataset was set aside during hyperparameter tuning and model training. These hyperparameters were then used on the test set for evaluation of the model's generalization performance (Appendix 4). To compare the performance of regression models and thus choose the optimal one for further analysis, root mean square error (RMSE) and ratio of performance to interquartile range (RPIQ) values were computed by applying the trained models with the optimized hyperparameters on the test set. Generally, a model with good prediction performance would have larger values of RPIQ and smaller values of RMSE.

Table 4. 2 The tuned hyperparameters and their ranges for each regression model.

| Regression Model | Hyperparameter | Range |
|---|---|---|
| Elastic net | Constant that multiplies the penalty terms | 0.01, 0.1, 1, 10, 100 |
| | Mixing parameter | 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 |
| Random forest regression | The number of variables to be considered for the best split | "auto", "sqrt", "log2" |
| | The maximum depth of the tree | 2 |
| | The number of trees in the forest | 100 |
| Support vector regression | The used kernel type | "linear", "poly", "rbf" |
| | Kernel coefficient | "scale", "auto" |
| | Regularization parameter | 0.01, 0.1, 1, 10, 100 |
| | The width of the epsilon-tube | 0.1, 0.5, 0.9 |

Notes: "Auto" refers to the total number of variables, "sqrt" refers to the square root of the total number of variables, "log2" refers to the binary logarithm of the total number of variables, "poly" refers to polynomial, "rbf" refers to radial basis function, "scale" refers to the use of 1/(total number of the variable × variance of the variables)

as the kernel coefficient, and "auto" refers to the use of 1/(total number of variable) as the kernel coefficient.

*4.2.8 Shapley Additive exPlanations Analysis*

The optimal model then underwent SHAP, based on game theory, to explore the relationships and quantify the contribution (SHAP values) of each input according to its average contribution to the model output (Lundberg & Lee, 2017). In this study, SHAP values based on KernelExplainer were computed for all samples since support vector regression performed the best at modeling and was used in the SHAP analysis. Summary plots for the whole dataset were generated to show important features and the directionality of their impact.

## 4.3 Results

*4.3.1 Variation in Stem Water Potential*

Both vineyards were visited nine times over two growing seasons, from flowering in late November to veraison in late January. This period is the most critical one before harvest in terms of the effects of GWS on berry quality. Figure 4.5 displays the variability in $\Psi_{stem}$ collected from 148 canopies at two vineyards and the distribution of the measurements on each date. The maximum and minimum observation of $\Psi_{stem}$ is 1344 and 293 kPa, respectively. Irrigation was only applied at the end of the study periods in the first growing season (Figure 4.2). Overall, there is an increasing trend of dehydration in GWS with time in both growing seasons, which indicates the impact of water deficit gradually accumulating in the canopies. The only exception is an increase in hydration state on 9 December 2021 compared to the previous measurement due to rainfall events a few days before sampling (i.e., 38.8 mm on 6 December 2021 and 11 mm on 7 December 2021). The height of the box and the difference between the upper quartile and lower quartile represent the spatial variation of $\Psi_{stem}$ on one date across each vineyard. Figure 4.5 demonstrates that it is inappropriate to make irrigation decisions during this period based on single or average measurements collected in the vineyard. This difference became larger as the survey proceeded, which implies spatial variation becomes more obvious when canopies are more dehydrated.

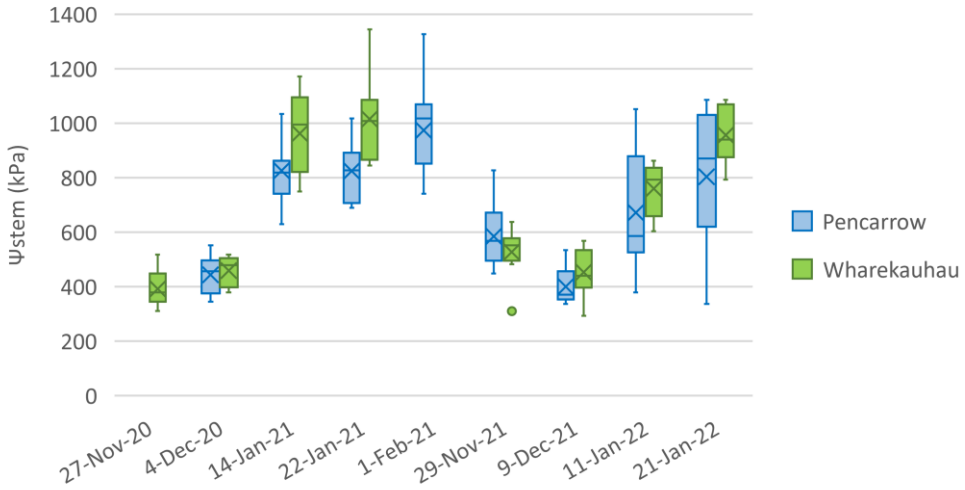



Figure 4. 5 Boxplot of measured stem water potential ($\Psi_{stem}$) values for the full set of samples collected at Pencarrow ($n$ = 86) and Wharekauhau ($n$ = 62) vineyard. X symbols refer to the average values on the survey dates. Lines in the boxes refer to median

values on the survey dates.

*4.3.2. Determination of the Best Descriptor of Vegetation Index for Variation in Grapevine Water Status*

According to the goal stated in the introduction section, 18 vegetation indices (VIs) frequently used in viticulture were computed using five bands (i.e., blue, green, red, red edge, and near-infrared) provided by the UAV-based multispectral sensors. VIs were compared with the corresponding $\Psi_{stem}$ values recorded in the field surveys to assess their correlation, in terms of $R^2$ and RMSE, and to check whether changes in $\Psi_{stem}$ can be reasonably assessed using drone-based imagery. Amongst all the indices, TCARI had the best linear correlation with variation in $\Psi_{stem}$, with a $R^2$ of 0.35 and a RMSE of 213kPa (Table 4.3). This was followed by ExG, with a $R^2$ of 0.3 and a RMSE of 221 kPa. Since TCARI is derived from green, red, and red edge bands, this suggests that TCARI is a promising candidate to help characterize the spatial variation of $\Psi_{stem}$ when multispectral sensors are employed. The promising performance of ExG implies that, when only RGB sensors are available, ExG may serve as a spectral indicator for $\Psi_{stem}$. Subsequently, these two indices were used as the core input for separate modeling, with weather, soil/terrain, and temporal variables used to explore the advantage of adding these environmental variables when regressing against $\Psi_{stem}$.

Table 4. 3 $R^2$ and RMSE based on a linear regression between $\Psi_{stem}$ and each vegetation index. The vegetation indices are ranked in descending order of $R^2$. r is Pearson correlation coefficient. $R^2$ is coefficient of determination, and RMSE is root mean square error. Abbreviations of Vegetation indices are described in full in Table 4.1.

| Vegetation Index | r | $R^2$ | RMSE (kPa) |
|---|---|---|---|
| TCARI | -0.59* | 0.35 | 213 |
| ExG | -0.55* | 0.30 | 221 |
| NDRE | -0.49* | 0.25 | 228 |
| TCARI/OSAVI | 0.49* | 0.24 | 231 |
| GNDVI | 0.50* | 0.24 | 231 |
| CL red edge | 0.48* | 0.24 | 231 |
| Canopy volume | 0.44* | 0.19 | 237 |
| MTVI1 | -0.39* | 0.16 | 243 |
| EVI | -0.35* | 0.13 | 247 |
| DVI | -0.34* | 0.12 | 248 |
| MSAVI | -0.31* | 0.10 | 251 |
| SR | 0.28* | 0.08 | 254 |
| NDVI | 0.21 | 0.04 | 258 |
| OSAVI | -0.18 | 0.03 | 260 |
| NGRDI | 0.13 | 0.02 | 262 |
| R/G index | -0.13 | 0.02 | 262 |
| VARI | -0.096 | 0.009 | 263 |
| MCARI | 0.013 | 0.0002 | 264 |

Significance levels are noted as * when $p \leq 0.001$

*4.3.3 Selection of Predictor Variables as Inputs for Modeling Using Hierarchical Clustering*

Hierarchical clustering, using correlation distance for distinguishing similarity, was employed to cluster the predictors that were statistically redundant to one another. Later, the model could use any variable from a cluster to form the input dataset. In Figure 4.6, each leaf of the dendrogram represents one of the 15 predictor variables. The vertical axis indicates how similar the variables are. The earlier (closer to the leaves) that merging between variables occurs, the more positively correlated those variables are. When variables merge close to the top of the dendrogram, the information they contain is nearly independent of each other. The number of clusters was determined by incorporating domain knowledge. As weather variables are intercorrelated and have an impact on plant physiology leading to changes in spectral features (i.e., VIs), it is reasonable to put them into clusters. However, the relationship between soil/terrain variables and short-term weather effects or VIs would be considered weak, so the cut-off line was set at 0.75 for both TCARI and ExG-based models, resulting in eight clusters marked with different colors. The summary for environmental variables used in this study was listed in Table 4.4.

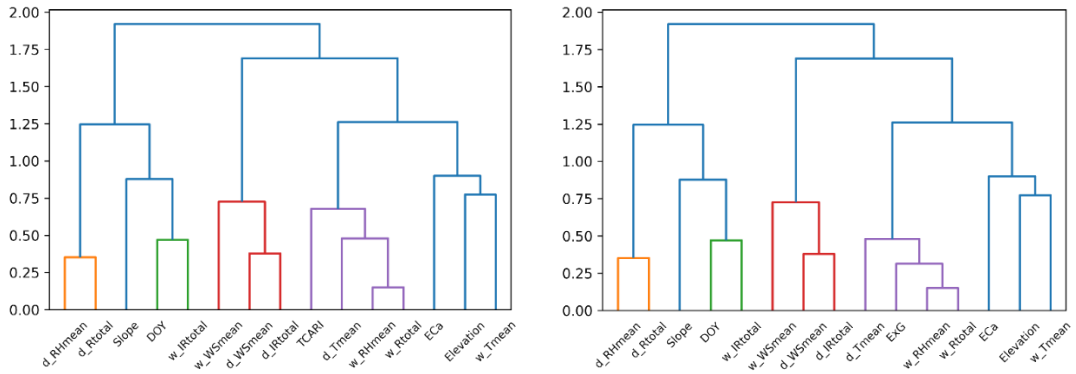



Figure 4. 6 Hierarchical clustering dendrogram of predictor variables: RHmean is mean relative humidity; Rtotal is total rainfall; DOY is day of the year; IRtotal is total irradiance; WSmean is mean wind speed; Tmean is mean temperature; $EC_a$ is apparent electrical conductivity. The initial letter (w or d) refers to the temporal scale as weekly or daily, respectively, before the measurement of stem water potential. (**Left**): predictor variables that use TCARI as a core input. (**Right**): predictor variables that use ExG as a core input.

Table 4. 4 A summary for all the environmental variables used to regress against changes in grapevine water status.

| **Predictor** | **Abbreviation or Short Name of Predictor** | **Type** |
|---|---|---|
| Day of the year | DOY | Temporal |
| Apparent electrical conductivity | $EC_a$ | Soil/terrain |
| Elevation | - | Soil/terrain |
| Slope | - | Soil/terrain |
| Mean relative humidity | RHmean | Weather |
| Total rainfall | Rtotal | Weather |
| Total irradiance | IRtotal | Weather |
| Mean wind speed | WSmean | Weather |

| | | |
|---|---|---|
| Mean temperature | Tmean | Weather |

*4.3.4 Regression of Grapevine Water Status Based on Core and Environmental Variables*

To test the potential usage of including the environmental variables, core variables (i.e., TCARI or ExG) and different combinations of the environmental variables composed the inputs for regression against $\Psi_{stem}$ using EN, RFR, or SVR. The core variable must always be a modeling component. Environmental variables were formed by one variable per cluster generated by hierarchical clustering, while the joining of variables from each cluster was optional. The total number of combinations tested was 576 (8 clusters: 3×2×3×4×1×2×2 ×2).

Compared to regression using VI only, there is a significant improvement in RMSE when including environmental variables as inputs (Table 4.3, the first rows in Tables 4.5 and 4.6). RMSE is improved from 213 to 146 kPa for the TCARI-based model, while RMSE is improved from 221 to 138 kPa for the ExG-based model. In both models, SVR, a non-linear algorithm, was chosen to be the best descriptor for the relationship between $\Psi_{stem}$ and the predictors. The scatter plots for both TCARI-based and ExG-based SVR models are presented in Figure 4.7.

From the second to the seventh rows in Tables 4.5 and 4.6, the modeling performance is presented to further assess the importance of environmental variables in terms of their type (temporal, soil/terrain, and weather). According to the types of inputs used, only the ML model with the best predictive performance among the three ML models, in terms of RMSE on the test set, is presented. For both TCARI-based and ExG-based models, those used soil/terrain and temporal variables as inputs to capture the most variance. In terms of one type of environmental predictor, models using terrain/soil variables as inputs have the worst performance for both TCARI-based and ExG-based models.

In terms of selected ML algorithms, both RFR and SVR are good descriptors to capture the variance of $\Psi_{stem}$ when regressed against two types of environmental variables (the second to fourth rows in Tables 4.5 and 4.6). RFR performs the best, shown by the fifth to seventh rows in Tables 4.5 and 4.6, for models based on one type of environmental predictor, and the results of both SVR and EN are not presented because only the models with the best performance, in terms of RMSE on the test set, are displayed based on the types of inputs used.

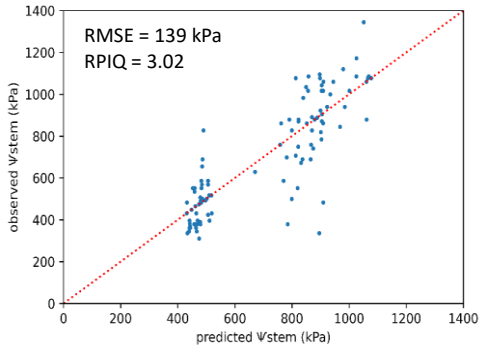
Table 4. 5 Regression modeling using TCARI as core input along with different types of environmental variables. RMSE is root mean square error, and RPIQ is ratio of performance to interquartile range.

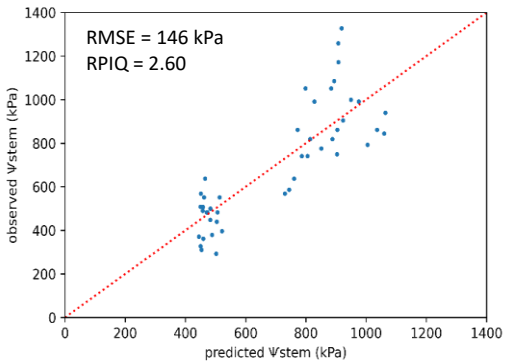
| Variable Composition | Machine Learning Algorithm | RMSE of the Train Set (kPa) | RPIQ of the Train Set | RMSE of the Test Set (kPa) | RPIQ of the Test Set |
|---|---|---|---|---|---|
| TCARI + full set of predictors | Support vector regression | 139 | 3.02 | 146 | 2.60 |
| TCARI + soil/terrain + weather | Random forest regression | 159 | 2.64 | 163 | 2.33 |
| TCARI + soil/terrain + temporal | Support vector regression | 139 | 3.02 | 146 | 2.60 |
| TCARI + weather + temporal | Random forest regression | 149 | 2.81 | 153 | 2.48 |

| | | | | | |
|---|---|---|---|---|---|
| TCARI + soil/terrain | Random forest regression | 165 | 2.54 | 172 | 2.21 |
| TCARI + temporal | Random forest regression | 149 | 2.81 | 150 | 2.53 |
| TCARI + weather | Random forest regression | 158 | 2.65 | 163 | 2.33 |

Table 4. 6 Regression modeling using ExG as core input along with different types of environmental variables. RMSE is root mean square error, and RPIQ is ratio of performance to interquartile range.

| Variable Composition | Machine Learning Algorithm | RMSE of the Train Set (kPa) | RPIQ of the Train Set | RMSE of the Test Set (kPa) | RPIQ of the Test Set |
|---|---|---|---|---|---|
| ExG + full set of predictors | Support vector regression | 134 | 3.14 | 138 | 2.75 |
| ExG + soil/terrain + weather | Random forest regression | 139 | 3.02 | 142 | 2.66 |
| ExG + soil/terrain + temporal | Support vector regression | 135 | 3.10 | 141 | 2.69 |
| ExG + weather + temporal | Support vector regression | 138 | 3.03 | 143 | 2.65 |
| ExG + soil/terrain | Random forest regression | 196 | 2.14 | 214 | 1.77 |
| ExG + temporal | Random forest regression | 128 | 3.27 | 159 | 2.39 |
| ExG + weather | Random forest regression | 138 | 3.04 | 144 | 2.64 |



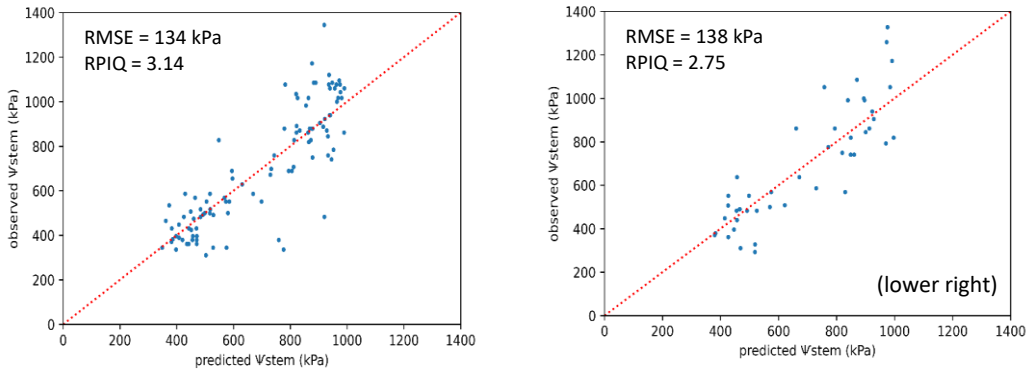



Figure 4. 7 Scatter plots between predicted stem water potential and observed stem water potential for TCARI-based training set (**upper left**), TCARI-based test set (**upper right**), ExG-based training set (**lower left**), and ExG-based test set using support vector regression based on the full set of predictors.

*4.3.5 Interpreting Models Using Shapley Additive exPlanations Analysis*

To explore the impact of each predictor variable on modeling output and variation in $\Psi_{stem}$, SHAP was deployed to support the interpretability of ML models. The SHAP values indicate the contribution of the variable towards model prediction. For Tables 4.5 and 4.6, the causal effects between $\Psi_{stem}$ and predictor variables were assessed using SHAP, and only those models with reasonable relationships were selected and presented.

Only models based on full types of predictors in Tables 4.5 and 4.6 were analyzed and presented using SHAP analysis. In the summary plots of Figures 4.8 and 4.9, every point from each predictor refers to an observation. A wider spread of points along the axis of SHAP values indicates the variable has more influence on the output, thus being of more importance. The color of the points denotes the value of the variable for the observation, where red and blue colors refer to the high and low variable values, respectively. The summary plot shows how each variable influences the model prediction. For example, a red point with a negative SHAP value indicates that a higher value of the variable will have a negative contribution to the $\Psi_{stem}$ value, the canopy becoming more hydrated. The bar plots of Figures 4.8 and 4.9 provide an overview of variable importance for each model. Both TCARI and ExG exhibit higher values when the grapevines are more hydrated. For the ExG-based model, it shows that higher wind speed leads grapevines to become more hydrated. The potential reasons underlying these relationships will be discussed in the next section.

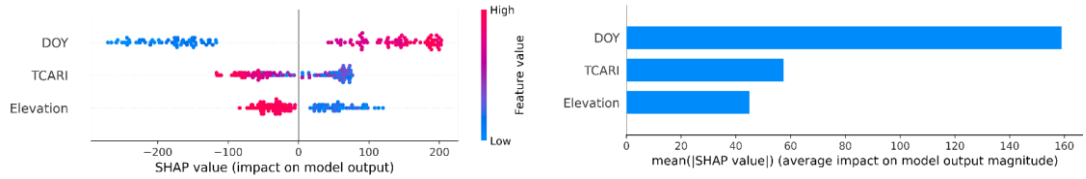



Figure 4. 8 (**Left**): summary plot for regression modeling using TCARI and selected environmental variables as inputs. (**Right**): bar plot for the predictor variables according to SHAP values received. Variables are ranked from top to bottom in descending order of average SHAP value, where DOY is day of the year; TCARI is transformed chlorophyll absorption reflectance index.

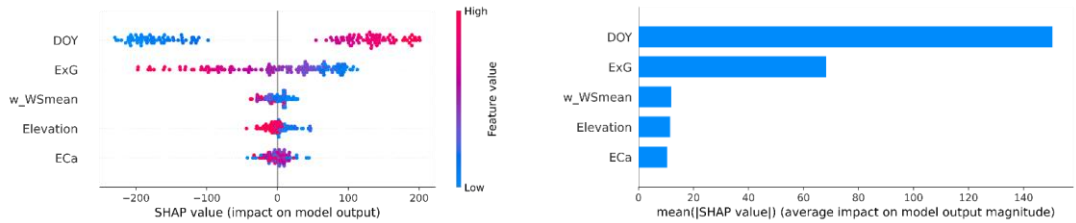



Figure 4. 9 (**Left**): summary plot for regression modeling using ExG and selected environmental variables as inputs. (**Right**): bar plot for the predictor variables according to SHAP values received. Variables are ranked from top to bottom in descending order of average SHAP value, where DOY is day of the year; ExG is excess green index; w_WSmean is weekly mean wind speed; $EC_a$ is apparent electrical conductivity.

## 4.4 Discussion

### *4.4.1 Vegetation Indices and Stem Water Potential*

Under water deficit, plants employ several adaptive strategies at leaf level, such as changes in leaf pigments and canopy structure, resulting in leaf curls and abscission (Ballester et al., 2019; Chaves et al., 2002; Turner & Begg, 1981). These pigment and structural changes cause variation in multispectral reflectance, including the blue, green, red, red edge, and near-infrared (NIR) bands. Multispectral reflectance detects water status indirectly via reflectance changes in response to physiological and morphological variability as a consequence of water stress (Poblete et al., 2017). The blue band is characterized by the absorption of chlorophyll and carotenoids, which can be used as an indicator of the resistance of plants to stress, such as water deficit (Gitelson, 2011; Steele et al., 2009). The energy in the green band is associated with the absorption by anthocyanins and reflection by chlorophyll (Viña et al., 2011). The red band has been reported to be associated with the concentration of chlorophyll (Ballester et al., 2018). The reflectance and position of the red edge band, a transition region between red and NIR bands, has been proven to be sensitive to chlorophyll content in leaves (Campbell et al., 2007; Clevers et al., 2002). The variation level of leaf water content alters cell turgidity and changes cell structure, which subsequently influences the absorption and reflection in the NIR band (Satterwhite & Henley, 1990). When crops experience water deficit, leaf chlorophyll content usually reduces, leading to a decrease in green reflection and an increase in blue and red reflection (Zarco-Tejada et al., 2000). In addition, there is an increase in the concentration of de-epoxidized components of xanthophyll (Chaves et al., 2002), while the alteration of canopy geometry results in a decrease in NIR reflection (Cogato et al., 2021). These basic characteristics lay the foundation for the use of vegetation indices (VIs) acquired from multispectral drones to monitor crop hydration status, as most of the current UAV systems do not include shortwave infrared bands (Maimaitijiang et al., 2020).

Most of the variability of $\Psi_{stem}$ in this study lies in non-stressed ($\Psi_{stem} \leq 800$ kPa) and moderately stressed (800 kPa $\leq \Psi_{stem} \leq$ 1200 kPa) ranges, while only three observations belong to the severely stressed range ($\Psi_{stem} \geq 1200$ kPa) (Acevedo-Opazo et al., 2008). Using single multispectral VI, the capability of capturing the variance in $\Psi_{stem}$ is poor, though the correlation is significant, with the best performance by TCARI, having an $R^2$ of 0.35 (Table 4.3). This unsatisfactory result is similar to that reported by Romero et al. (2018), in which the best performance of all VIs was an $R^2$ of 0.42. A possible explanation for this may be related to the sensitivity of canopy structure and leaf pigments to various levels of water stress for different crop species

(Ballester et al., 2018). For example, chlorophyll is affected after severe dehydration treatment or after an extended period of water deficit (Zulini et al., 2005), so VIs associated with chlorophyll variation may not be sensitive enough to detect mild water stress or short-term dehydration. Moreover, a study has shown that differences in spectral resolution could influence values of VIs, such as NDVI (Teillet et al., 1997), which can compromise the relationship between spectral indices and the variable of interest.

According to Zhang et al. (2021), multispectral VIs related to water deficit can be classified as chlorophyll and structural VIs based on their sensitivity to changes in physiological or morphological characteristics. Chlorophyll VIs are derived from blue, green, red, or red edge bands, so they are more sensitive to changes in chlorophyll concentration. TCARI, ExG, and TCARI/OSAVI belong to this category. Structural VIs are computed based on spectral bands, including NIR, so they are both influenced by chlorophyll and leaf scattering caused by changes in plant structure and involve the use of GNDVI, NDRE, and NDVI. It is interesting to note that chlorophyll VIs are relatively good at capturing the variance of $\Psi_{stem}$ (Table 4.3). This implies that the water stress level of the grapevine used in this study, Pinot Noir, is described better via observation of chlorophyll-related VIs compared to structure-related VIs. Research carried out at Pinot Noir vineyards in New Zealand over two growing seasons (Mejias-Barrera, 2016) stated that dehydration treatment did not have an overall effect on the reduction of leaf area and that this effect started becoming evident from veraison onwards. They highlighted the usefulness of using chlorophyll content in water stress studies, which supports our results. The top five VIs (TCARI, ExG, TCARI/OSAVI, GNDVI, NDRE, and one frequently used VI, NDVI), in terms of its $R^2$, are the focus for the following discussion.

Among the chlorophyll VIs, TCARI ($R^2$ = 0.35) explains more variance in $\Psi_{stem}$ than ExG ($R^2$ = 0.3) does in terms of their correlation coefficients, followed by TCARI/OSAVI ($R^2$ = 0.24). TCARI is a VI, which is sensitive to chlorophyll concentration, while TCARI/OSAVI has been proposed to minimize the effects posed by soil reflectance and changes in leaf area index (LAI) (Haboudane et al., 2002). The better performance of TCARI compared to the modified index (TCARI/OSAVI) may indicate that the LAI of grapevines was relatively stable and high, thus swamping any noise from the soil background during the study period. This finding is supported by the study of Baluja et al. (2012), in which TCARI ($R^2$ = 0.45) and TCARI/OSAVI ($R^2$ = 0.58) were significantly correlated with the $\Psi_{stem}$ of rain-fed Tempranillo grapevines in Spain. TCARI yields a better correlation with $\Psi_{stem}$ than ExG, probably due to the inclusion of the red edge band. This behavior was also observed at Petite Sirah and Cabernet Sauvignon vineyards in California, USA (Tang et al., 2022). ExG is a VI highlighting a green band to separate vegetation from the background, and thus is sensitive to canopy greenness (Woebbecke et al., 1995). A good correlation between ExG and GWS was also found to be evident for grapevines faced with different irrigation treatments at a Vermentino vineyard in Italy (Matese et al., 2016). The good performance of ExG suggests that it has the potential of detecting GWS using just RGB sensors instead of multispectral sensors. This may increase the applicability of UAV-based imagery because RGB cameras do not require an extra filter and post-processing work to generate NIR readings, thus reducing the cost of sensors (Hunt et al., 2010).

Both GNDVI and NDRE have been reported to be sensitive to variations in chlorophyll concentration (Boiarskii & Hasegawa, 2019; Gitelson & Merzlyak, 1998). Due to the modification of chlorophyll content under water stress, GNDVI was observed to be able to differentiate Cabernet Sauvignon grapevines between the

moderate and severe classes of water deficit (Espinoza et al., 2017). A good correlation between GNDVI and $\Psi_{stem}$ was recorded by Baluja et al. (2012) and Poblete et al. (2017) for Carménère grapevines planted in Chile, in which $R^2$ was 0.58 and 0.31, respectively. NDRE was observed to outperform other multispectral indices when correlated with leaf water potential, with a Pearson correlation coefficient of −0.27 (Tang et al., 2022). NDVI is the most used VI in agriculture, as it is related to plant biomass, vigor, and health. In viticulture, the growth of shoots is inhibited after experiencing water deficit (Schultz & Matthews, 1988), and that development state, quantified by pruning weight and LAI, has been shown to be strongly related to NDVI (Caruso et al., 2017). Acevedo-Opazo et al. (2008) have observed a significant difference in pre-dawn leaf water potential between high and low NDVI zones for seven cultivars grown in France. However, in this study, the correlation between NDVI and $\Psi_{stem}$ is weak ($R^2$ = 0.04). The poor performance of NDVI is supported by Romero et al. (2018), whose research indicated that the red and NIR bands could be used as long-term predictors of water deficit as correlation improved over the season. Another explanation is the saturation effect of NDVI values, which usually happens when applied to very dense canopies. In the research of Junges et al. (2017), a lower standard deviation of 0.01 was calculated when NDVI was measured close to 0.8 for Cabernet Sauvignon grapevines planted in Brazil. This value is similar to this study's NDVI recordings (average = 0.74; maximum = 0.82; standard deviation = 0.06).

### *4.4.2 Important Environmental Variables: Day of the Year, Elevation, Electrical Conductivity, and Wind Speed*

Inspired by the concept of a soil-plant-atmosphere continuum for water movement in nature (Elfving et al., 1972), various ancillary estimators (other than VIs) have been used to help describe the dynamics of GWS (Table 4.4). Plant water status during the day is a consequence of water uptake (mainly dependent on characteristics of the root system and soil moisture), as well as transpiration (mainly dependent on a number of atmospheric variables and stomatal conductance) (Rodríguez et al., 2010). In this study, it was assumed that $\Psi_{stem}$ was more strongly associated with soil water status and weather conditions. This assumption was made because the cultivar effect has been removed, as only Pinot Noir with similar canopy ages were examined. In addition, daylight is assumed to be the main driver to stomatal opening in this study. As the temperature is mild and rainfall is moderate in Martinborough, the occurrence of additional processes such as stomatal closure induced by abscisic acid resulting from soil water deficit may be reduced. Moreover, Pinot Noir is considered to be a near-anisohydric cultivar, so its control on stomata is less responsive to decreases in soil moisture (Gutiérrez-Gamboa et al., 2019).

The selected core VIs (TCARI and ExG) exhibiting chlorophyll variability are the variables representing the mid-term impact imposed by water deficit, as the chlorophyll content changes in accordance with the level of dehydration during the growing period. It is assumed that the predictive capacities for $\Psi_{stem}$ can be significantly improved when complementing VIs with comparable long-term and short-term variables. Soil/terrain variables represent long-term effects, as they consistently affect GWS over years (Bellvert et al., 2012). The temporal variable is relatively mid- to short-term because it reflects the time trend along the growing season. Weekly or daily weather variables are short term. The complementary effects among different types of variables are reflected in Tables 4.5 and 4.6. When using two types of environmental predictors, the composition of predictors (VI, soil/terrain, and

temporal variables) captures the most variance in $\Psi_{stem}$ in both TCARI-based and ExG-based models. When using solely one type of environmental predictor, the combination of VI and temporal or weather variables produces models with the best performance, similar to the performance of the models using two types of environmental predictors. This indicates that using variables with mid- to short-term effects may be sufficient to describe the variation of $\Psi_{stem}$ in this study. TCARI and ExG-based models had the worst performance when combined with soil/terrain variables, while the ExG-based model performed worse than the TCARI-based model. This may be because the ExG captures relatively mid to long-term effects of water stress on chlorophyll content, while TCARI captures relatively short to mid-term effects. This supports the result when using a full set of environmental predictors (Figures 4.8 and 4.9). The best performing TCARI-based model does not select any weather variables, and the ExG-based model selects one weather variable as an input.

The summary plots in Figures 4.8 and 4.9 provide insights into the relationships between $\Psi_{stem}$ and the selected environmental variables (DOY), elevation, $EC_a$, and weekly mean wind speed. DOY represents seasonality along the growing season. DOY positively contributes to the dehydration of the grapevine (Figures 4.8 and 4.9). Irrigation was only applied at the end of the study periods in the first growing season, with heavy precipitation concentrated around flowering and fruit set and subsequent gentle (<10 mm) rainfall (Figure 4.2). As a result, there is a downward trend of $\Psi_{stem}$ (increasing hydration) from late November to late January (Figure 4.5) in accordance with the increase in DOY. In the study of Suter et al. (2019), DOY explained a great part of the variability in the $\Psi_{stem}$ of three cultivars grown in France.

SHAP summary plots show that grapevines are more hydrated when they stand on more elevated ground. It is contrary to the concept in which water accumulation happens in the lower parts of fields so that crops would stay at high hydration levels. When evaluating terrain information of observations at both vineyards, Pencarrow, ranging between 44–46 m elevation, is higher than Wharekauhau, ranging between 36–38 m (Appendix 2). The recorded elevation of observations from Wharekauhau is significantly lower than that of observations from Pencarrow (*p*-value of Mann-Whitney U test is lower than 0.05). One explanation is that the higher landscape terraces, older than 15K years, are likely to have a covering of loess. Loess is a well-sorted quartzo-feldspathic material of silt to clay-sized particles, and it is picked up by wind from nearby aggrading braided riverbeds present during historic global cold cycles. This fine loess will increase the water holding capacity of the soils developing on these higher and older terraces. Loess deposition will not be present on lower terraces with an age younger than 15K years, where the soils are much more likely to be developing from alluvial deposits from consecutive flooding events over the last 15K years. These deposits have a potentially wider range of textures from silt to sand and coarse sand. The soils developing from them are likely to be better drained (Molly & Hewitt, 1988). The soil properties were reported to affect soil water holding capacity, further leading to spatial differences in GWS across the fields (Brillante et al., 2016). Thus, Pencarrow, sited on the higher terrace, may have the ability to hold more water than Wharekauhau due to differences in particle composition, enabling grapevines to be more hydrated.

$EC_a$ was selected as an important predictor for the ExG-based model. $EC_a$ has been demonstrated to be related to spatial changes in GWS ($R^2 = 0.56$) due to its association with differences in soil water availability (Yu et al., 2020). The contribution of EC to $\Psi_{stem}$ (Figure 4.9) does not follow a regular pattern. Higher clay content usually leads to higher $EC_a$ because clay particles can adsorb more water (Lal & Shukla, 2004).

However, the $EC_a$ survey in this study was taken during a wet period, and soil $EC_a$ recorded in this kind of period was influenced by both soil moisture content and clay content (Zhu et al., 2010). Therefore, the relationship between $EC_a$ and $\Psi_{stem}$ is not straightforward. Other sources of variation include the spatial sensitivity of the EM38 instrument (Callegary et al., 2012), the size of the root ball of each canopy, and the physical properties of the soils hosting the functional roots. These factors will determine how precisely $EC_a$ values represent the soil properties influencing each grapevine.

Of the weather variables, only weekly mean wind speed (mean = 1.5 m/s, standard deviation = 0.6 m/s, and maximum = 3 m/s) was selected by the ExG-based model (Figure 4.9). Crop transpiration is influenced by weather through the effects on the driving force, including VPD and the resistance, including stomatal resistance ($r_s$) and boundary layer resistance ($r_b$). VPD is a function of temperature, which depends on solar radiation and relative humidity. Solar radiation, temperature, and wind speed impact $r_s$ and $r_b$, with wind speed dominating (Keller, 2020). In the summary plots (Figures 4.8 and 4.9), it is interesting to note that there is a general trend of negative response of grapevine dehydration to windspeed, as blue points positively contribute to $\Psi_{stem}$. This is contradictory to the general understanding in which water stress is induced by both increasing evapotranspiration demands and decreasing boundary layers, both resulting from high wind speed (Jarvis & McNaughton, 1986). These results are supported by several studies that have observed grapevines reducing transpiration, via partial closure of stomata, to prevent water loss. Two studies measured this phenomenon when wind speeds are higher than 3.6 m/s for six cultivars grown in pots (Kobriger et al., 1984) and 4 m/s for two cultivars planted in Western Australia (Campbell-Clause, 1998). An additional reason could be more efficient convective cooling, resulting from a thinner boundary layer or high humidity of the air circulating around leaves, removing heat from the irradiated leaf surface, rather than relying on evaporative cooling (Schymanski & Or, 2016). Temperature, irradiance, humidity, and rainfall are widely accepted as important weather factors influencing the variation in $\Psi_{stem}$. The reason those variables were not selected in our models was due to the length of the study window. In this study, the two months between flowering to veraison were selected due to their significance to GWS monitoring. This may not be sufficient to sample broader variability in these weather variables.

### *4.4.3 Regression Modeling*

The relationship between $\Psi_{stem}$ and soil/terrain, temporal, and weather variables is complex, so ML algorithms, including EN, RFR, and SVR, were utilized to describe the relationship. As $R^2$ is an inadequate measure for nonlinear models (Spiess & Neumeyer, 2010), it was not used in evaluating modeling performance. According to the results listed in Tables 4.5 and 4.6, the relationship seems to be non-linear, since EN is a linear regression model with penalty terms and does not outperform RFR or SVR when using any set of predictors. When modeling is based on all types of predictors, SVR has the best performance in both TCARI-based and ExG-based models. When based on two types of predictors, either SVR or RFR is selected as the best descriptor. When only one type of predictor is used in regression, RFR outperforms the other algorithms. It appears SVR is capable of handling more variables, especially when soil/terrain and temporal variables are included in simulating changes in $\Psi_{stem}$. One explanation is that this dataset may be limited ($n$ = 148) and have appreciable noise. However, SVR has been reported to be less sensitive to sampling variation for a

small sample size because the selection of support vectors depends on only a small subset of observations (James et al., 2013; Tange et al., 2017). Moreover, the close values between RMSE for training and test sets show the overfitting issue in our models was addressed well. This indicates that the range of the regularized hyperparameters (the maximum depth of the tree for RFR, regularization parameter, and width of the epsilon-tube for SVR) are set correctly for the study. Interpretation of the results supports using SVR for remote monitoring of Pinot Noir $\Psi_{stem}$ if the regression is based on terrain/soil, temporal, and weather variables, but more ML algorithms need to be investigated to optimize the estimation performance.

### *4.4.4 Limitations and Future Work*

By combining UAV-based multispectral imagery with ancillary variables, the approach presented in this study demonstrates the potential usefulness of using UAVs to estimate GWS at canopy scale across the vineyard when the locations of perennial grapevines in a vineyard are precisely marked. However, this study uses an empirical approach, so it does not fully describe the relationships between soil/terrain, weather, spectral information, and grapevine water status. As summarized by Adams et al. (2013), empirical approaches are highly dependent on the range and the quality of inputs and can yield significant bias when used to extrapolate identified relationships beyond observed variability. The methodology demonstrated in this study could serve as a useful tool under the following conditions: the training datasets incorporate a rich variation in high-quality values of predictors corresponding to a range of sites, environmental conditions, grapevine cultivars, growing stages, canopy ages, and management practices. ML modeling is undertaken with extensive calibration. Additional variables, such as soil moisture, can be investigated for their usefulness as ancillary variables. To enhance the modeling performance and transferability, it will be necessary to enlarge the input datasets. This is a time-consuming process, involving the collection of a large number of $\Psi_{stem}$ measurements by pressure bomb within the two hours of midday. This will require considerable manpower and calibrated pressure bombs. Once these conditions are met, the model developed in this study may be applied to other vineyards, cultivars, or, perhaps, regional scale analysis.

## 4.5 Conclusions

This study explores the potential of combining UAV-based vegetation indices (VIs) and various environmental data (soil/terrain, temporal, and weather variables) for the estimation of stem water potential ($\Psi_{stem}$), used as a proxy for grapevine water status (GWS). It uses machine learning algorithms and SHAP analysis, based on observations collected in two Pinot Noir vineyards in New Zealand, over two growing seasons. The results demonstrate the potential and the techniques associated with developing a useful tool for $\Psi_{stem}$ monitoring by taking advantage of the complementary effects between spectral, soil/terrain, weather, and temporal variables. SHAP analysis facilitates the interpretability of the $\Psi_{stem}$ model, which could be useful for GWS manipulation. A broader range of predictor values is needed to train and calibrate the models for enhancing the transferability of GWS models to other vineyards. In addition, it is recommended that new machine learning algorithms be investigated to optimize the performance of $\Psi_{stem}$ estimation. For viticulturists and growers, the tool developed in this study could enable precise scheduling of irrigation for quality management from canopy to sub-block scale in response to the types of irrigation equipment available.

**References**


Acevedo-Opazo, C., Tisseyre, B., Guillaume, S., & Ojeda, H. (2008). The potential of high spatial resolution information to define within-vineyard zones related to vine water status. *Precision Agriculture, 9*(5), 285-302. https://doi.org/10.1007/s11119-008-9073-1

Acevedo-Opazo, C., Tisseyre, B., Taylor, J., Ojeda, H., & Guillaume, S. (2010). A model for the spatial prediction of water status in vines (Vitis vinifera L.) using high resolution ancillary information. *Precision Agriculture, 11*(4), 358-378.

Acevedo-Opazo, C., Valdés-Gómez, H., Taylor, J., Avalo, A., Verdugo-Vásquez, N., Araya, M., Jara-Rojas, F., & Tisseyre, B. (2013). Assessment of an empirical spatial prediction model of vine water status for irrigation management in a grapevine field. *Agricultural Water Management, 124*, 58-68.

Adams, H. D., Williams, A. P., Xu, C., Rauscher, S. A., Jiang, X., & McDowell, N. G. (2013). Empirical and process-based approaches to climate-induced forest mortality models. *Frontiers in Plant Science, 4*, 438.

Arevalo-Ramirez, T., Villacrés, J., Fuentes, A., Reszka, P., & Cheein, F. A. A. (2020). Moisture content estimation of Pinus radiata and Eucalyptus globulus from reconstructed leaf reflectance in the SWIR region. *Biosystems Engineering, 193*, 187-205.

Baciocco, K. A., Davis, R. E., & Jones, G. V. (2014). Climate and Bordeaux wine quality: identifying the key factors that differentiate vintages based on consensus rankings. *Journal of Wine Research, 25*(2), 75-90.

Ballester, C., Brinkhoff, J., Quayle, W. C., & Hornbuckle, J. (2019). Monitoring the effects of water stress in cotton using the green red vegetation index and red edge ratio. *Remote Sensing, 11*(7), 873.

Ballester, C., Zarco-Tejada, P., Nicolás, E., Alarcón, J. J., Fereres, E., Intrigliolo, D. S., & Gonzalez-Dugo, V. (2018). Evaluating the performance of xanthophyll, chlorophyll and structure-sensitive spectral indices to detect water stress in five fruit tree species. *Precision Agriculture, 19*(1), 178-193.

Ballesteros, R., Ortega, J. F., Hernández, D., & Moreno, M. Á. (2015). Characterization of Vitis vinifera L. canopy using unmanned aerial vehicle-based remote sensing and photogrammetry techniques. *American Journal of Enology and Viticulture, 66*(2), 120-129.

Baluja, J., Diago, M. P., Balda, P., Zorer, R., Meggio, F., Morales, F., & Tardaguila, J. (2012). Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV). *Irrigation Science, 30*(6), 511-522. https://doi.org/10.1007/s00271-012-0382-9

Barnes, E., Clarke, T., Richards, S., Colaizzi, P., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., & Thompson, T. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, USA,

Bellvert, J., Marsal, J., Mata, M., & Girona, J. (2012). Identifying irrigation zones across a 7.5-ha 'Pinot noir'vineyard based on the variability of vine water status and multispectral images. *Irrigation Science, 30*(6), 499-509.

Birth, G. S., & McVey, G. R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer 1. *Agronomy Journal, 60*(6), 640-643.

Boiarskii, B., & Hasegawa, H. (2019). Comparison of NDVI and NDRE indices to detect differences in vegetation and chlorophyll content. *Journal of mechanics of continua and mathematical sciences, 4*, 20-29.

Bowyer, P., & Danson, F. (2004). Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. *Remote Sensing of Environment, 92*(3), 297-308.

Bramley, R., Ouzman, J., & Boss, P. K. (2011). Variation in vine vigour, grape yield and vineyard soils and topography as indicators of variation in the chemical composition of grapes, wine and wine sensory attributes. *Australian Journal of Grape and Wine Research, 17*(2), 217-229.

Brevik, E. C., Fenton, T. E., & Lazari, A. (2006). Soil electrical conductivity as a function of soil water content and implications for soil mapping. *Precision Agriculture, 7*(6), 393-404.

Brillante, L., Martínez-Luscher, J., Yu, R., Plank, C. M., Sanchez, L., Bates, T. L., Brenneman, C., Oberholster, A., & Kurtural, S. K. (2017). Assessing spatial variability of grape skin flavonoids at the vineyard scale based on plant water status mapping. *Journal of Agricultural and Food Chemistry, 65*(26), 5255-5265.

Brillante, L., Mathieu, O., Lévêque, J., & Bois, B. (2016). Ecophysiological modeling of grapevine water stress in burgundy terroirs by a machine-learning approach. *Frontiers in Plant Science, 7*, 796.

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.

Callegary, J. B., Ferré, T. P., & Groom, R. (2012). Three-dimensional sensitivity distribution and sample volume of low-induction-number electromagnetic-induction instruments. *Soil Science Society of America Journal, 76*(1), 85-91.

Campbell-Clause, J. (1998). Stomatal response of grapevines to wind. *Australian Journal of Experimental Agriculture, 38*(1), 77-82.

Campbell, P. K., Middleton, E. M., McMurtrey, J. E., Corp, L. A., & Chappelle, E. W. (2007, May-Jun). Assessment of vegetation stress using reflectance or fluorescence measurements. *J Environ Qual, 36*(3), 832-845. https://doi.org/10.2134/jeq2005.0396

Caruso, G., Tozzini, L., Rallo, G., Primicerio, J., Moriondo, M., Palai, G., & Gucci, R. (2017). Estimating biophysical and geometrical parameters of grapevine canopies ('Sangiovese') by an unmanned aerial vehicle (UAV) and VIS-NIR cameras. *Vitis, 56*(2), 63-70.

Chaves, M. M., Pereira, J. S., Maroco, J., Rodrigues, M. L., Ricardo, C. P., Osório, M. L., Carvalho, I., Faria, T., & Pinheiro, C. (2002). How plants cope with water stress in the field? Photosynthesis and growth. *Annals of Botany, 89*(7), 907-916.

Choné, X. (2001). Stem Water Potential is a Sensitive Indicator of Grapevine Water Status. *Annals of Botany, 87*(4), 477-483. https://doi.org/10.1006/anbo.2000.1361

Clevers, J., De Jong, S., Epema, G., Van Der Meer, F., Bakker, W., Skidmore, A., & Scholte, K. (2002). Derivation of the red edge index using the MERIS standard band setting. *International Journal of Remote Sensing, 23*(16), 3169-3184.

Cogato, A., Wu, L., Jewan, S. Y. Y., Meggio, F., Marinello, F., Sozzi, M., & Pagay, V. (2021). Evaluating the Spectral and Physiological Responses of Grapevines (Vitis vinifera L.) to Heat and Water Stresses under Different Vineyard Cooling and Irrigation Strategies. *Agronomy, 11*(10), 1940.

Cook, P. G., & Williams, B. G. (1998). *Electromagnetic Induction Techniques-Part 8*. CSIRO PUBLISHING.

Daughtry, C. S., Walthall, C., Kim, M., De Colstoun, E. B., & McMurtrey Iii, J. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment, 74*(2), 229-239.

Elfving, D. C., KAUFMANN, M. R., & HALL, A. E. (1972). Interpreting leaf water potential measurements with a model of the soil-plant-atmosphere continuum. *Physiologia Plantarum, 27*(2), 161-168.

Espinoza, C. Z., Khot, L. R., Sankaran, S., & Jacoby, P. W. (2017). High resolution multispectral and thermal remote sensing-based water stress assessment in subsurface irrigated grapevines. *Remote Sensing, 9*(9), 961.

Etchebarne, F., Ojeda, H., & Hunter, J. (2010). Leaf: fruit ratio and vine water status

effects on Grenache Noir (Vitis vinifera L.) berry composition: water, sugar, organic acids and cations. *South African Journal of Enology and Viticulture, 31*(2), 106-115.

Gamon, J., & Surfus, J. (1999). Assessing leaf pigment content and activity with a reflectometer. *The New Phytologist, 143*(1), 105-117.

Giovos, R., Tassopoulos, D., Kalivas, D., Lougkos, N., & Priovolou, A. (2021). Remote sensing vegetation indices in viticulture: A critical review. *Agriculture, 11*(5), 457.

Gitelson, A. (2011). Nondestructive estimation of foliar pigment (chlorophylls, carotenoids and anthocyanins) contents: Evaluating a semianalytical three-band model. In P. S. Thenkabail, & Lyon, J.G. (Ed.), *Hyperspectral Remote Sensing of Vegetation* (pp. 141-166). CRC Press. https://doi.org/10.1201/b11222

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment, 80*(1), 76-87.

Gitelson, A. A., & Merzlyak, M. N. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research, 22*(5), 689-692.

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters, 32*(8).

Gutiérrez-Gamboa, G., Pérez-Donoso, A. G., Pou-Mir, A., Acevedo-Opazo, C., & Valdés-Góme, H. (2019). Hydric behaviour and gas exchange in different grapevine varieties (Vitis vinifera L.) from the Maule Valley (Chile). *South African Journal of Enology and Viticulture, 40*(2), 1-1.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment, 90*(3), 337-352.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment, 81*(2-3), 416-426.

Heil, K., & Schmidhalter, U. (2017). The application of EM38: Determination of soil parameters, selection of soil sampling points and use in agriculture and archaeology. *Sensors, 17*(11), 2540.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*(1-2), 195-213.

Hunt, E. R., Hively, W. D., Fujikawa, S. J., Linden, D. S., Daughtry, C. S., & McCarty, G. W. (2010). Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sensing, 2*(1), 290-305.

Intrigliolo, D. S., & Castel, J. R. (2009). Response of grapevine cv. 'Tempranillo' to timing and amount of irrigation: water relations, vine growth, yield and berry and wine composition. *Irrigation Science, 28*(2), 113-125. https://doi.org/10.1007/s00271-009-0164-1

Irmak, S., & Mutiibwa, D. (2010). On the dynamics of canopy resistance: Generalized linear estimation and relationships with primary micrometeorological variables. *Water resources research, 46*(8).

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jang, G., Kim, J., Yu, J.-K., Kim, H.-J., Kim, Y., Kim, D.-W., Kim, K.-H., Lee, C. W., & Chung, Y. S. (2020). Cost-effective unmanned aerial vehicle (UAV) platform for field plant breeding application. *Remote Sensing, 12*(6), 998.

Jarvis, P. G., & McNaughton, K. G. (1986). Stomatal control of transpiration: scaling up from leaf to region. In *Advances in ecological research* (Vol. 15, pp. 1-49). Elsevier.

Jenal, A., Bareth, G., Bolten, A., Kneer, C., Weber, I., & Bongartz, J. (2019). Development of a VNIR/SWIR multispectral imaging system for vegetation monitoring with unmanned aerial vehicles. *Sensors, 19*(24), 5507.

Junges, A. H., Fontana, D. C., Anzanello[1], R., & Bremm, C. (2017). Normalized difference vegetation index obtained by ground-based remote sensing to characterize vine cycle in Rio Grande do Sul, Brazil. *Ciência e Agrotecnologia, 41*, 543-553.

Kandylakis, Z., Falagas, A., Karakizi, C., & Karantzalos, K. (2020). Water Stress Estimation in Vineyards from Aerial SWIR and multispectral UAV data. *Remote Sensing, 12*(15), 2499.

Keller, M. (2020). *The science of grapevines*. Academic press.

Kobriger, J., Kliewer, W., & Lagier, S. (1984). Effects of wind on water relations of several grapevine cultivars. *American Journal of Enology and Viticulture, 35*(3), 164-169.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer.

Lal, R., & Shukla, M. K. (2004). *Principles of soil physics*. CRC Press.

Liu, C., SUN, P.-S., & Liu, S.-R. (2016). A review of plant spectral reflectance response to water physiological changes. *Chinese Journal of Plant Ecology, 40*(1), 80.

López-García, P., Intrigliolo, D. S., Moreno, M. A., Martínez-Moreno, A., Ortega, J. F., Pérez-Álvarez, E. P., & Ballesteros, R. (2021). Assessment of Vineyard Water Status by Multispectral and RGB Imagery Obtained from an Unmanned Aerial Vehicle. *American Journal of Enology and Viticulture, 72*(4), 285-297.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence, 2*(1), 56-67.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Proceedings of the 31st international conference on neural information processing systems,

Maimaitijiang, M., Sagan, V., Sidike, P., Daloye, A. M., Erkbol, H., & Fritschi, F. B. (2020). Crop monitoring using satellite/UAV data fusion and machine learning. *Remote Sensing, 12*(9), 1357.

Mangalathu, S., Hwang, S.-H., & Jeon, J.-S. (2020). Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Engineering Structures, 219*, 110927.

Martínez-Lüscher, J., Sánchez-Díaz, M., Delrot, S., Aguirreolea, J., Pascual, I., & Gomes, E. (2014). Ultraviolet-B radiation and water deficit interact to alter flavonol and anthocyanin profiles in grapevine berries through transcriptomic regulation. *Plant and Cell Physiology, 55*(11), 1925-1936.

Matese, A., Baraldi, R., Berton, A., Cesaraccio, C., Di Gennaro, S., Duce, P., Facini, O., Mameli, M., Piga, A., & Zaldei, A. (2016). Combination of proximal and remote sensing methods for mapping water stress conditions of grapevine. International Symposium on Sensing Plant Water Status-Methods and Applications in Horticultural Science 1197,

Mejias-Barrera, P. (2016). *Effect of reduced irrigation on grapevine physiology, grape characteristics and wine composition in three Pinot noir vineyards with contrasting soils* Lincoln University].

Min, Z., Li, R., Chen, L., Zhang, Y., Li, Z., Liu, M., Ju, Y., & Fang, Y. (2019). Alleviation of drought stress in grapevine by foliar-applied strigolactones. *Plant Physiology and Biochemistry, 135*, 99-110.

Molloy, L., & Hewitt, A. E. (1988). *Soils in the New Zealand landscape: the living mantle* (pp. 179-181). Wellington: Mallison Rendel.

Morgenthaler, S. (2009). Exploratory data analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 1*(1), 33-44.

Ojeda, H., Andary, C., Kraeva, E., Carbonneau, A., & Deloire, A. (2002). Influence of pre- and postveraison water deficit on synthesis and concentration of skin phenolic compounds during berry growth of Vitis vinifera cv. Shiraz. *American Journal of Enology and Viticulture, 53*(4), 261-267.

Patakas, A., Noitsakis, B., & Chouzouri, A. (2005). Optimization of irrigation water use in grapevines using the relationship between transpiration and plant water status. *Agriculture, Ecosystems & Environment, 106*(2-3), 253-259.

Poblete, T., Ortega-Farias, S., Moreno, M. A., & Bardeen, M. (2017, Oct 30). Artificial Neural Network to Predict Vine Water Status Spatial Variability Using Multispectral Information Obtained from an Unmanned Aerial Vehicle (UAV). *Sensors (Basel), 17*(11). https://doi.org/10.3390/s17112488

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment, 48*(2), 119-126.

Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing, 109*, 88-97. https://doi.org/10.1016/j.isprsjprs.2015.09.003

Rodríguez, J. C., Grageda, J., Watts, C. J., Garatuza-Payan, J., Castellanos-Villegas, A., Rodríguez-Casas, J., Saiz-Hernandez, J., & Olavarrieta, V. (2010). Water use by perennial crops in the lower Sonora watershed. *Journal of Arid Environments, 74*(5), 603-610. https://doi.org/10.1016/j.jaridenv.2009.11.008

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. https://doi.org/10.1016/j.compag.2018.02.013

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55*(2), 95-107.

Rouse, J., Haas, R., Schell, J., Deering, D., & Harlan, J. (1974). Monitoring the vernal advancements and retrogradation. *Texas AM Univ, Texas*.

Satterwhite, M. B., & Henley, J. P. (1990). *Hyperspectral signatures (400 to 2500 nm) of vegetation, minerals, soils, rocks, and cultural features: Laboratory and field*

*measurements*. A. E. T. L. F. B. VA.

Schultz, H. R., & Matthews, M. A. (1988). Vegetative growth distribution during water deficits in Vitis vinifera L. *Functional Plant Biology, 15*(5), 641-656.

Schymanski, S. J., & Or, D. (2016). Wind increases leaf water use efficiency. *Plant, cell & environment, 39*(7), 1448-1459.

Spiess, A.-N., & Neumeyer, N. (2010). An evaluation of $R^2$ as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC pharmacology, 10*(1), 1-11.

Steele, M. R., Gitelson, A. A., Rundquist, D. C., & Merzlyak, M. N. (2009). Nondestructive estimation of anthocyanin content in grapevine leaves. *American Journal of Enology and Viticulture, 60*(1), 87-92.

Suter, B., Triolo, R., Pernet, D., Dai, Z., & Van Leeuwen, C. (2019). Modelling stem water potential by separating the effects of soil water availability and climatic conditions on water status in grapevine (Vitis vinifera L.). *Frontiers in Plant Science, 10*, 1485.

Tang, Z., Jin, Y., Alsina, M. M., McElrone, A. J., Bambach, N., & Kustas, W. P. (2022). Vine water status mapping with multispectral UAV imagery and machine learning. *Irrigation Science*, 1-16.

Tange, R. I., Rasmussen, M. A., Taira, E., & Bro, R. (2017). Benchmarking support vector regression against partial least squares regression and artificial neural network: Effect of sample size on model performance. *Journal of Near Infrared Spectroscopy, 25*(6), 381-390.

Taylor, J. A., Acevedo-Opazo, C., Pellegrino, A., Ojeda, H., & Tisseyre, B. (2012). Can within-season grapevine predawn leaf water potentials be predicted from meteorological data in non-irrigated Mediterranean vineyards? *OENO One, 46*(3), 221-232.

Taylor, J. A., Acevedo-Opazo, C., Ojeda, H., & Tisseyre, B. (2010). Identification and significance of sources of spatial variation in grapevine water status. *Australian Journal of Grape and Wine Research, 16*(1), 218-226.

Teillet, P., Staenz, K., & William, D. (1997). Effects of spectral, spatial, and radiometric characteristics on remote sensing vegetation indices of forested regions. *Remote Sensing of Environment, 61*(1), 139-149.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*(2), 127-150.

Turner, N. C., & Begg, J. E. (1981). Plant-water relations and adaptation to stress. *Plant*

*and Soil, 58*(1), 97-131.

Van Leeuwen, C., Trégoat, O., Choné, X., Bois, B., Pernet, D., & Gaudillère, J.-P. (2009). Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? *OENO One, 43*(3), 121-134.

Viña, A., Gitelson, A. A., Nguy-Robertson, A. L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment, 115*(12), 3468-3478.

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE, 38*(1), 259-269.

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*.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., & Noland, T. L. (2000). Chlorophyll fluorescence effects on vegetation apparent reflectance: I. Leaf-level measurements and model simulation. *Remote Sensing of Environment, 74*(3), 582-595.

Zhang, L., Han, W., Niu, Y., Chávez, J. L., Shao, G., & Zhang, H. (2021). Evaluating the sensitivity of water stressed maize chlorophyll and structure based on UAV derived vegetation indices. *Computers and Electronics in Agriculture, 185*, 106174.

Zhu, Q., Lin, H., & Doolittle, J. (2010). Repeated electromagnetic induction surveys for determining subsurface hydrologic dynamics in an agricultural landscape. *Soil Science Society of America Journal, 74*(5), 1750-1762.

Zulini, L., Rubinigg, M., Zorer, R., & Bertamini, M. (2005). Effects of drought stress on chlorophyll fluorescence and photosynthetic pigments in grapevine leaves (Vitis vinifera cv.'White Riesling'). International Workshop on Advances in Grapevine and Wine Research 754.

Chapter 4 studies the ancillary effects of various environmental variables in estimating changes in GWS. The results show that there is a potential to employ the complementary abilities these environmental variables offer to develop a GWS monitoring tool using multispectral sensors. To carry out large-scale monitoring, spaceborne platforms serve as a feasible solution for GWS tracking. Taking advantage of this chapter's findings allows grape growers to estimate the spatial variability of GWS in the fields using multispectral satellites. Nevertheless, relatively low resolution

for satellite images and the delay in data acquisition by end users hinder the usability of applying multispectral satellites to viticultural management. Thus, the next chapter is aimed to solve these two issues and enhance the applicability of multispectral spaceborne sensors in GWS management.

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: | Hsiang-En Wei |
| Name and title of main supervisor: | Dr Miles Grafton, Senior Lecturer |
| In which chapter is the manuscript/published work? | Chapter 5 |
| What percentage of the manuscript/published work was contributed by the student? | 82.5% |

Describe the contribution that the student has made to the manuscript/published work:

The candidate used two-stage calibration to improve the accuracy of grapevine water status estimation using satellite images while developing a grapevine water status prediction model/tool for assisting irrigation management.

Please select one of the following three options:

| | |
|---|---|
| ○ | **The manuscript/published work is published or in press**<br>Please provide the full reference of the research output: |
| ◉ | **The manuscript is currently under review for publication**<br>Please provide the name of the journal:<br>Technology in Agronomy (accepted for publication) |
| ○ | **It is intended that the manuscript will be published, but it has not yet been submitted to a journal** |

| | | | |
|---|---|---|---|
| Student's signature: | Hsiang-En Wei 數位簽署者：Hsiang-En Wei 日期：2023.04.11 16:46:10 +12'00' | 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: 2023.05.04 21:15:55 +12'00' |

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

# Chapter 5
# Evaluation of the use of Two-Stage Calibrated PlanetScope Images and Environmental Variables for the Development of the Grapevine Water Status Prediction Model

This chapter is dedicated to evaluating the third hypothesis – The reliable prediction of GWS at a large scale can be achieved by the application of geospatial processing and modeling using ML algorithms. The objective is developing a prediction tool based on multispectral satellite images that can provide the estimation of GWS at a large scale in advance. The complementary effects provided by the environmental variables, from soil/terrain, temporal, weather, and vegetation perspectives, for characterizing spatial GWS are shown. These effects were utilized to facilitate multispectral spaceborne sensors to provide a larger-scale GWS monitoring. It is focused on addressing two issues when applying satellite images to vineyard management, by utilizing two-level spatial calibration and establishing prediction models. Their performance is validated by applying the prediction model to a dataset collected in an independent year.

**Abstract:** Grapevine water status (GWS) assessment between flowering and veraison plays an important role in viticulture management in terms of producing high-quality grapes. Although satellites and uncrewed aerial vehicles (UAV) have successfully monitored GWS, these platforms are practically limited because data transfer is delayed due to post processing and UAV operation is weather dependent. This study focuses on addressing two issues: the unreliability of GWS estimation using satellite images with low-moderate spatial resolution and the inaccessibility of real-time satellite data. It aims to predict the temporal variation of GWS based on a prediction model using spectral information (calibrated PlanetScope (PS) images), soil/topography data (apparent electrical conductivity, elevation, slope), weather parameters (rainfall and potential evapotranspiration), cultivation practices (irrigation, fertigation, plucking, and trimming), and seasonality (day of the year) as predictors. Stem water potential ($\Psi_{stem}$) was used as a proxy for GWS. Two-stage calibration, including an initial calibration of UAV images with measured $\Psi_{stem}$ and a subsequent calibration of satellite images with calibrated UAV data, was applied to calibrate the PS images. Three machine learning models (random forest regression, support vector regression, and multilayer perceptron) were used in the calibration and modeling process. The results showed that a two-stage calibration can generate reliable reference data, with a root mean square error of 113 kPa and 59 kPa on the test sets during the first and second calibration stage, respectively. The prediction model described the temporal variation of block $\Psi_{stem}$ well when compared with the measured $\Psi_{stem}$. In the similarity analysis, the Pearson correlation coefficient was 0.89 and 0.87 between predicted and reference $\Psi_{stem}$ maps across four dates for the two studied vineyards. This study supports the concept of developing an approach to predict grapevine $\Psi_{stem}$, which would enable growers to acquire $\Psi_{stem}$ variation in advance during the growing season, leading to improved irrigation scheduling and optimal grape quality.

## 5.1 Introduction

Plant water status is widely recognized as an important factor in attaining a high-quality grape due to its role in vegetative growth, partitioning mechanism, berry development, and metabolic composition (Chaves et al., 2010; Martínez-Lüscher et al., 2014; Ojeda et al., 2002; Van Leeuwen et al., 2009). However, grapevine water status (GWS) exhibits spatial variability across vineyards under a homogeneous irrigation regime, leading to variation in berry development and quality (Bramley et al., 2011; Brillante et al., 2017). Management of the spatial variability of GWS across vineyards can potentially benefit growers due to berry quality optimization (Arnó Satorra et al., 2009; Monaghan et al., 2013). To help achieve optimization of quality, provision of spatial (across vineyards) and temporal (along growing season) GWS monitoring maps would be very useful to the vineyard managers.

Electromagnetic reflectance obtained from remote sensing (RS) platforms, such as satellites and uncrewed aerial vehicles (UAVs), has gained popularity for acting as proxies for GWS (Matese et al., 2015; Sozzi et al., 2020). RS techniques enable spatial monitoring of GWS within vineyards, so its application can be applied to develop site-

specific irrigation or assist in decision-making for irrigation management. Compared to satellites, UAVs mounted with sensors have the relative advantages of flexibility in flight scheduling, imagery with a higher spatial resolution (at centimeter-scale), and lower cost of operation. However, UAV surveys are confined to smaller areas due to the short flight endurance imposed by their payload (Matese et al., 2015). This may lead to the introduction of non-negligible uncertainty in the radiometric quality of orthoimages resulting from mosaicking a large number of tiles (Borgogno-Mondino, Lessio, et al., 2018). Satellites, compared to UAVs, are superior when mapping at a larger scale within a single acquisition on a regular basis, guaranteeing radiometric homogeneity for all pixels in the scene (Borgogno-Mondino, Lessio, et al., 2018). Nevertheless, when considering crops with discontinuous layouts, such as vineyards, satellite imagery with low to moderate resolution cannot be reliably used for plant status description due to the biased assessment originating from the presence of inter-row vegetation and bare soil (Helman et al., 2018; Khaliq et al., 2019). Although there are some satellites (e.g. GEOSAT-2, GeoEye-1, WorldView-3) that offer images with high resolution and can be used for field monitoring in viticulture, the cost of image acquisition makes it impractical in the application of irrigation scheduling that needs regular monitoring. The challenge when using RS tools for timely monitoring of highly dynamic phenomena, such as GWS monitoring, is to obtain data that not only has high spatial and temporal resolution but also is affordable and timely accessible.

Pla et al. (2019) used UAV data as reference to calibrate Sentinel-2 imagery to quantify damage in rice crops. This study demonstrated that the calibration approach can serve as a viable and cost-efficient alternative to retrieve biophysical variables at large scales. Revill et al. (2020) later developed a two-stage method for the calibration of satellite vegetation indices (VIs) to approximate crop leaf area index by calibrating UAV images with ground measurements, followed by calibrating satellite images with calibrated UAV data. This new methodological framework provides an opportunity to integrate the advantages of both satellites and UAVs by utilizing UAV data to help bridge ground truthing and satellite images. This will enable variables of interest to be retrieved at a larger scale, while removing the spectral response of the inter-row elements during the calibration process. This will enhance the reliability of the use of satellite data for plant status monitoring. Based on the study of Revill et al. (2020), Bukowiecki et al. (2021) transferred this approach to green area index estimation based on the data collected over four growing seasons. The satellite-based prediction model achieved a high performance (coefficient of determination of 0.82). Mazzia et al. (2020) used a convolutional neural network to refine satellite-based normalized difference VI maps by utilizing UAV-derived information. They found the refined maps could better describe crop status in terms of correlation analysis and ANOVA, in comparison to the raw satellite images. However, the delivery of satellite imagery to the end user is not in real-time after the fields are sensed. There are often delays in data transfer to the ground station, handling, and distribution to the user, while some processes (such as geometric, radiometric, and atmospheric corrections) are needed before image analysis (Gautam & Pagay, 2020). In addition, the application of satellite imagery over open fields is weather-dependent and illumination sensitive. Cloud and fog interference is a common restriction imposed on the usability of satellite-based data for crop monitoring. This requires the user to wait for the next revisit of data acquisition (Brook et al., 2020; Devaux et al., 2019). These constraints are likely to compromise the practicability of the satellite imagery for growers and viticulturists, despite the satellite images being calibrated with ground truthing. Thus, there is a

clear requirement to develop tools for GWS forecasting to promote precision irrigation and thus berry quality.

To potentially address the inaccessibility of satellite-based data caused by unfavorable weather, one way is to develop a GWS prediction model using a soil water balance. As GWS is an integrative response to soil moisture availability, water usage by plants, evaporation, atmospheric transfer, and canopy structures (van Leeuwen et al., 2006), there is a potential to simulate the changes in GWS along a growing season by using these variables as predictors. Precipitation, irrigation, and fertigation replenish soil water content, thus positively impacting GWS (Jasse et al., 2021). Evapotranspiration (quantifying the net loss of water vapor from evaporation and plant transpiration) accounts for the dynamics and interrelationships of weather components, vegetation expression, and soil properties (Allen et al., 1998). Soil properties and topography have persistent effects on the spatial distribution of soil water, which determines the soil moisture availability to plants, and thus may influence GWS (Bellvert et al., 2012). Cultivation management, such as leaf plucking and trimming, alters canopy vigor, shoot distribution, and total leaf area, leading to the regulation of water consumption level by grapevines (Ohana-Levi et al., 2020). Seasonality, noted as day of the year (DOY), has been found to explain a large part of the variability in GWS (Suter et al., 2019).

Machine learning (ML) models have recently been extensively used in RS applications due to their ability in modeling both linear and non-linear relationships (Kuhn & Johnson, 2013). These models can be applied to multi-dimensional problems, so they can potentially simulate spectral, spatial, and temporal variabilities between images from different platforms, as well as the complex relationships between GWS and predictors. Previous studies have shown the potential of using UAV or satellite-derived VIs to assess GWS (Baluja et al., 2012; Bellvert et al., 2012; Borgogno-Mondino, et al., 2018; Espinoza et al., 2017; Helman et al., 2018; Matese et al., 2018; Poblete et al., 2017; Romero et al., 2018). However, these assessments were mostly limited to identifying the relationships between UAV-based or satellite-based data, and GWS, without considering the need for real-time monitoring which is constrained by the time required for post-processing and associated weather-dependent issues. Most of the RS platforms are currently not able to provide real-time information which is especially important in the context of dynamic GWS monitoring. This study aims to provide GWS prediction based on two-stage calibrated satellite images, which caps off the two previous studies (Wei et al., 2021; Wei et al., 2022) with a prediction tool to assist growers and viticulturists in irrigation scheduling or related management decisions for vineyards. Wei et al. (2021) focused on identifying the most related spectral features with changes in GWS, and it turned out that they dispersed across multiple spectral bands and needed to be computed collectively using multivariate models. Wei et al. (2022) highlighted the importance of environmental variables (vegetation data, temporal trends, weather parameters, and soil/terrain characteristics) in helping a multispectral sensor describe the variation of GWS. The procedures of this study undertaken were (i) using two-stage calibration (based on ML models) to acquire calibrated satellite images exhibiting GWS estimations for a series of acquisition dates across the first growing season. (ii) developing a GWS prediction model by regressing against the calibrated satellite images acquired in the first growing season using ML models. (iii) evaluating this prediction model by validation with a second set of ground measurements independently collected in the second growing season. To the authors' knowledge, this is the first study that a two-stage calibration approach has been employed to calibrate satellite images for developing

GWS prediction model based on ancillary information collected in the first growing season, along with independent evaluation with data collected the second growing season.

## 5.2 Materials and methods

### *5.2.1 The Context of the Study Vineyards and Study periods*

The study vineyards are located at Martinborough in the Greater Wellington Region (lower North Island) in New Zealand (NZ). Both vineyards are sited on a complex of recent alluvial soils overlying gravels, located close to the Ruamahanga and Huangarua Rivers (Figure 5.1). The study site contains two commercial vineyards owned by Palliser Estate, named Wharekauhau and Pencarrow. Our study areas in these two vineyards are 6.6 and 6.7 ha, respectively. Pinot Noir was chosen as the target variety in this study, due to its requirement for relatively precise GWS management. The Pinot Noir vines in both vineyards were planted in the vineyards in 1998-2000 and two-cane vertical shoot positioned. Inter- and intra-row planting space is 2.2 × 1.6 m for Wharekauhau and 2.2 × 1.8 m for Pencarrow. The annual growth cycle of grapevine in NZ comprises budburst, shoot growth, and flowering (September–November), fruit set and veraison (December–February), followed by berry development and harvesting (March–May). From flowering to veraison, the management of GWS critically determines the final berry quality (Fernandes-Silva et al., 2019; Schreiner & Lee, 2014). The study periods span 15$^{th}$ November 2020 to 2$^{nd}$ February 2021 for the first growing season (2020/2021) and 23$^{rd}$ November 2021 to 21$^{st}$ January 2022 for the second growing season (2021/2022). This study focused on cultivation practices, including irrigation, fertigation, plucking, and trimming, due to their potential effects on GWS. Both vineyards are drip-irrigated, with the distance between drippers being 0.6 m. The drip rate is 1.2 L/hr for the drippers in Wharekauhau, and 1.3 L/hr for the drippers in Pencarrow. The irrigation events took place for 2.5 hours on each irrigation date. Fertigation was applied to provide additional nutrients, and lasted for 6 hours on each date of fertigation. Leaf plucking was implemented to remove 80% of the leaves at one side of the canopy to allow sprays better access to grapevines. Trimming was carried out on top of, and on both sides of, canopies to control vegetative growth. Weather information and cultivation practice events are displayed in Figure 5.2.


Figure 5. 1 Location of study vineyards.

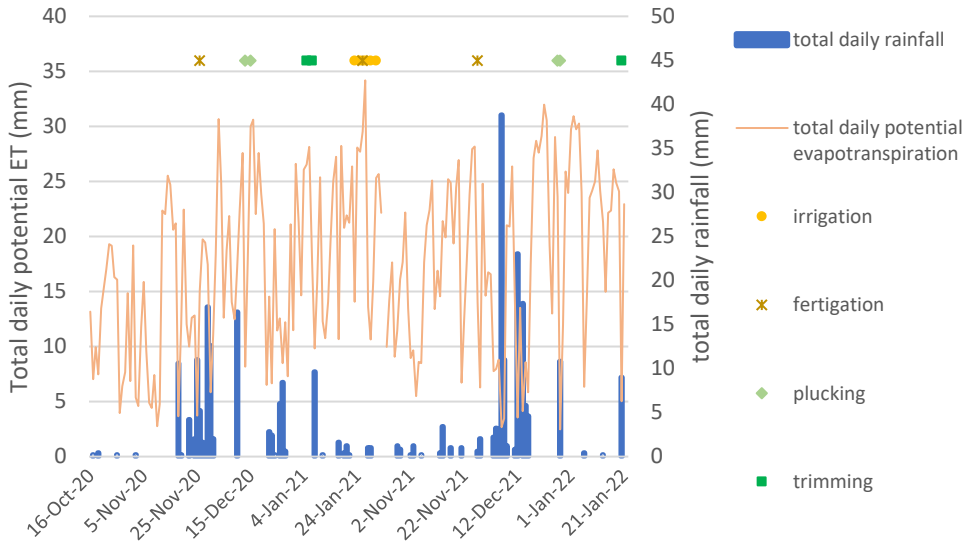



Figure 5. 2 Total daily potential evapotranspiration and total daily rainfall provided by an on-site weather station, with cultivation practices implemented during the two growing seasons.

### *5.2.2 Measured data*

Stem water potential ($\Psi_{stem}$) was chosen as an indicator for GWS because it has been reported as a comprehensive expression for early water deficit in vines during the day (Patakas et al., 2005). Several healthy vines were sampled in grids to assess variability across the vineyards for each measurement date (Table 5.1) using two mature and fully expanded leaves from the middle part of each sampled canopy.

These leaves are more representative of the canopies. A pressure chamber (model: 610, MPS, Albany, NY, USA) was used between 13:00 and 15:30 to acquire $\Psi_{stem}$ (kPa). Prior to the measurement, the sampled leaves were covered with sealable plastic bags for approximately 1 hour. The mean value of the two leaf measurements per vine represent the canopy $\Psi_{stem}$. A total of 85 and 63 separate canopies were sampled in the first and the second growing season, respectively (Appendix 1). Each of their trunk locations was recorded using a global navigation satellite system (GNSS) with real-time kinematic (RTK) correction (model: GPS1200+, Leica Geosystems AG., Heerbrugg, Switzerland).

Table 5. 1 Details of data acquisition for measured $\Psi_{stem}$, DJI Phantom 4 multispectral UAV imagery, and PlanetScope satellite imagery. The orange outlines indicate the UAV-satellite image pairs used in the second calibration. UAV is uncrewed aerial vehicle.

| Acquisition Date | Data Source | Time Gap (day) |
|---|---|---|
| 15/11/2020 | Satellite | - |
| 16/11/2020 | Satellite | - |
| 27/11/2020 | Measured/ UAV | - |
| 02/12/2020 | Satellite | - |
| 04/12/2020 | Measured/ UAV | 1 |
| 05/12/2020 | Satellite | |
| 15/12/2020 | Satellite | - |
| 17/12/2020 | Satellite | - |
| 31/12/2020 | Satellite | - |
| 04/01/2021 | Satellite | - |
| 14/01/2021 | Measured/ UAV/ Satellite | 0 |
| 15/01/2021 | Satellite | - |
| 22/01/2021 | Measured/ UAV | 1 |
| 23/01/2021 | Satellite | |
| 26/01/2021 | Satellite | - |
| 01/02/2021 | Measured/ UAV | 1 |
| 02/02/2021 | Satellite | |
| 23/11/2021 | UAV | - |
| 29/11/2021 | Measured/ UAV | - |
| 09/12/2021 | Measured/ UAV | - |
| 11/01/2022 | Measured/ UAV | - |
| 21/01/2022 | Measured/ UAV | - |

*5.2.3 UAV-based data*

UAV images were collected between 11:00 and 13:00 under cloudless conditions to minimize the impacts of sun angle and shadow. UAV flights and $\Psi_{stem}$ measurements were operated on the same dates to ensure comparability (Table 5.1; Appendix 8). The reflectance, with a spatial resolution of 0.05 m, was recorded by a DJI Phantom 4 multispectral UAV (DJI, Shenzhen, China) with six built-in sensors, including blue (450 ± 16 nm), green (560 ± 16 nm), red (650 ± 16 nm), red edge (730 ± 16 nm), and near-infrared (840 ± 26 nm) regions. An integrated sunlight sensor on the top of the UAV records irradiance in the same bands captured by the multispectral

sensor during the flight. This information enables UAV images to be normalized and, allows for comparison between images taken under different illumination conditions. Pix4Dmapper (Pix4D SA, Lausanne, Switzerland) was used to apply photogrammetric processing to the UAV data to generate digital surface models (DSM), digital terrain models (DTM), and reflectance maps. To increase imagery spatial accuracy, several ground control points were obtained by GNSS-RTK for each vineyard, and post-imagery alignment was performed in ArcGIS Pro 2.9 (ESRI, Redlands, California, USA).

The vineyards are characterized by discontinuous vegetation surfaces, requiring separation of canopy pixels from grass and soil pixels to obtain pure information of grapevines. As there is a height difference between grapevines and their surrounding components, canopy height can be acquired by subtracting DTM from DSM, then creating binary imagery to exclude background pixels. There are 16 vegetation indices, chosen according to the frequency of usage in viticulture (Giovos et al., 2021), shown in Table 5.2 and calculated based on pure canopy pixels using "zonal statistic as table" in ArcGIS Pro.

Table 5. 2 List of vegetation indices used in this study.

| **Vegetation Index** | **Formula** | **Reference** |
|---|---|---|
| Red Edge Chlorophyll Index | (NIR/Red edge) − 1 | (Gitelson et al., 2005) |
| Difference Vegetation Index | NIR − Red | (Tucker, 1979) |
| Enhanced Vegetation Index | 2.5 × (NIR − Red)/(NIR + 6 * Red − 7.5 × Blue + 1) | (Huete et al., 2002) |
| Excess Green Index | 2 × Green − Red − Blue | (Woebbecke et al., 1995) |
| Green Normalized Difference Vegetation Index | (NIR − Green)/(NIR + Green) | (Gitelson & Merzlyak, 1998) |
| Modified Chlorophyll Absorption Ratio Index | ((Red edge − Red) − 0.2 × (Red edge − Green)) × (Red edge/Red) | (Daughtry et al., 2000) |
| Modified Soil Adjusted Vegetation Index | $(2 \times NIR+1 - ((2 \times NIR + 1)^2 - 8 \times (NIR - Red))^{1/2})/2$ | (Qi et al., 1994) |
| Modified Triangular Vegetation Index | 1.2 × (1.2 × (NIR − Green) − 2.5 × (Red − Green)) | (Haboudane et al., 2004) |
| Normalized Difference Red Edge Index | (NIR − Red edge)/(NIR + Red edge) | (Barnes et al., 2000) |
| Normalized Difference Vegetation Index | (NIR − Red)/(NIR + Red) | (Rouse et al., 1974) |
| Normalized Difference Green/Red Index | (Green − Red)/(Green + Red) | (Tucker, 1979) |
| Optimized Soil Adjusted Vegetation Index | (NIR − RED)/(NIR + Red + 0.16) | (Rondeaux et al., 1996) |
| Red: Green Ratio | Red/Green | (Gamon & Surfus, 1999) |

| | | |
|---|---|---|
| Simple Ratio | NIR/Red | (Birth & McVey, 1968) |
| Transformed Chlorophyll Absorption Reflectance Index | 3 × ((Red edge − Red) − 0.2 × (Red edge − Green) × (Red edge/Red)) | (Haboudane et al., 2002) |
| Visible Atmospherically Resistant Index | (Green − Red)/(Green + Red − Blue) | (Gitelson et al., 2002) |

*5.2.4 Satellite-based data*

PlanetScope (PS) imagery was chosen for this study mainly due to its high spatial and temporal resolution (Appendix 8). PS Constellation, owned by Planet, is composed of approximately 200 "Dove" microsatellites, operating in sun-synchronous orbits and able to provide daily land surface imagery at nadir for the entire Earth, passing the equator between 9:30 - 11:30 (local solar time) (Planet Team, 2017). The product used in this study is PS Analytic Ortho Scene Surface Reflectance, which consists of level 3B images recorded by instrument PSB.SD in coastal blue (431-452 nm), blue (465-515 nm), green I (513-549 nm), green II (547-583 nm), yellow (600-620 nm), red (650 - 680 nm), red edge (697-713 nm), and near-infrared (845-885 nm) bands. These images are acquired with a ground sample distance of 3.7-4.1 m and then resampled to 3.0 m for distribution. This product is orthorectified to remove geometric distortions, and its positional accuracy is less than 10 m RMSE at the $90^{th}$ percentile. Radiometric corrections are applied to convert pixel values to at-sensor radiance, while atmospheric corrections are implemented to acquire surface reflectance values. Although PS imagery is superior in terms of spatiotemporal resolution, there is radiometric inconsistency due to cross-satellite differences in spectral response, radiometric quality, and orbital configuration (Houborg & McCabe, 2018; Leach et al., 2019). This heterogeneity was addressed in this study by incorporating the radiometric variability in the modeling processes. PS Images, covering the study vineyards and obtained on 13 different days over the study period in 2020/2021, were selected for further analysis after cloud and cloud shadow screening (Table 5.1).

*5.2.5 Soil and Terrain Data*

An EM38-MK2 is an electromagnetic induction (EMI)-based sensor (Geonics Ltd., Mississauga, Ontario, Canada). The return reading, apparent soil electrical conductivity ($EC_a$), is considered a function of soil texture, soil solution, and soil water content (Cook & Williams, 1998). $EC_a$ was reported to estimate the spatial variation of GWS (Yu et al., 2020). In this study, the EM38-MK2 was operated in the vertical dipole mode, with the instrument taking integrated $EC_a$ measurements at about 1.5 m depth. An EMI survey was undertaken by towing the EM38-MK2 at the back of an all-terrain vehicle, with a Trimble Yuma tablet (including an onboard GPS receiver) to geo-reference all $EC_a$ points (mS/m). The total rainfall one week before the survey is 7 mm, and two weeks before the survey is 15.8 mm. The vineyards' subsurface infrastructure was confirmed with the vineyard manager to ensure there was no interference from buried metal objects. $EC_a$ points were measured approximately every 3-10 m along transects that were 10 m apart. Point values less than 0 mS/m were removed before interpolation. The geostatistical interpolation method, Empirical Bayesian Kriging, was used to convert point data into a raster with 1 m resolution (Appendix 3).

Elevation (m) for the vineyards was obtained from the 'Wellington LiDAR 1m DEM (2013-2014)' layer provided by Land Information New Zealand data service (https://data.linz.govt.nz/). This digital elevation model (in 1 m resolution) was created by aerial LiDAR for the Greater Wellington region (captured between 2013 and 2014) (Appendix 2). Slope (degree) information was derived based on elevation.

### *5.2.6 Weather and Temporal Data*

Weather data was recorded by the on-site weather station (175.4741, -41.2247 WGS84) established by HARVEST.com (http://harvest.com/). The variables captured include rainfall (mm) and potential evapotranspiration (PET; mm). Rainfall was assumed to be homogeneous across the two vineyards. PET, based on the recordings collected by the weather station, was provided by HARVEST.com. PET represents the energy-driven water demand for ET by a short green crop (Penman, 1948). The temporal dependence of GWS on dates through the season was represented by DOY.

### *5.2.7 Regression Modeling*

Random forest regression (RFR), support vector regression (SVR), and multilayer perceptron (MLP) were employed for the image calibration model and $\Psi_{stem}$ prediction model in this study. They were implemented using "RandomForestRegressor", "SVR" and "MLPRegressor" from the sklearn library in Python 3.9. As the performance of models is influenced by their hyperparameters, hyperparameter tuning was undertaken beforehand to prevent overfitting. This enabled these ML models to exploit their potential. Grid searching on the training set with 10-fold cross-validation was used to search for the best set of hyperparameters. The test dataset was set aside during hyperparameter tuning and model training. These hyperparameters would then be used on the test set for evaluation of the model's generalization performance (Appendix 4). This technique was carried out using "GridSearchCV" from the sklearn library in Python. For each modeling in the two-stage calibration and prediction model, samples were split as training and test sets using a 70/30 ratio. The split was implemented and stratified according to the date of measurement, to ensure that both training and test sets had corresponding percentages of samples for each date. The splitting process was undertaken using "train_test_split" from the sklearn package in Python. All predictor variables were standardized to have mean values equivalent to 0 and a standard deviation of 1 before modeling.

### *5.2.8 Evaluation*

To compare the performance of the image calibration models and the $\Psi_{stem}$ prediction models, root mean square error (RMSE) values were computed by applying the trained models with optimized hyperparameters to the test set. The models with the lowest RMSE were chosen for further analysis. To measure the similarity of $\Psi_{stem}$ between predicted maps (obtained from the $\Psi_{stem}$ prediction model) and the reference maps (acquired in 2021/2022), comparisons were conducted cell by cell using Pearson correlation. The closer the value of the correlation coefficient (r) to ±1, the stronger the linear relationship. This correlation was implemented using "pearsonr" from the scipy library in Python.

### *5.2.9 Processing Workflow*

To develop the $\Psi_{stem}$ prediction model, two-stage calibration modeling was carried out (Figure 5.3). It was assumed the measured $\Psi_{stem}$ data accounts for spatial and temporal variation of $\Psi_{stem}$ across the two vineyards over the study periods, and the relationships between $\Psi_{stem}$ and spectral data are stable. First, 0.05m resolution UAV images were calibrated using measured $\Psi_{stem}$ data collected during 2020/2021 (n = 85) and 2021/2022 (n = 63), respectively (Figure 5.3 (A)). The calibrated 0.05m UAV images (Figure 5.3 (B)) were used as reference images for the calibration of satellite images and the validation of the $\Psi_{stem}$ prediction model. The calibrated 0.05m UAV images acquired during 2020/2021 (n = 4), representing $\Psi_{stem}$ estimation, were subsequently downscaled to 3m resolution for spatial matching with the satellite images (Figure 5.3 (C)), by spatially averaging the pixel values in the corresponding cells of the PS 3m grids. Next, these downscaled 3m UAV images (n = 4) were used to calibrate the PS images obtained in 2020/2021 (Figure 5.3 (E)). All calibrated PS images (n = 13; Figure 5.3 (F)) obtained in the study period during 2020/2021 were then used to develop the $\Psi_{stem}$ prediction model (Figure 5.3 (G)). The developed model (Figure 5.3 (H)) was then validated by the downscaled 3m UAV images (n = 5) acquired during 2021/2022 (Figure 5.3 (I)). These were generated by spatially aggregating the calibrated 0.05m UAV images acquired in 2021/2022 (Figure 5.3 (D)).

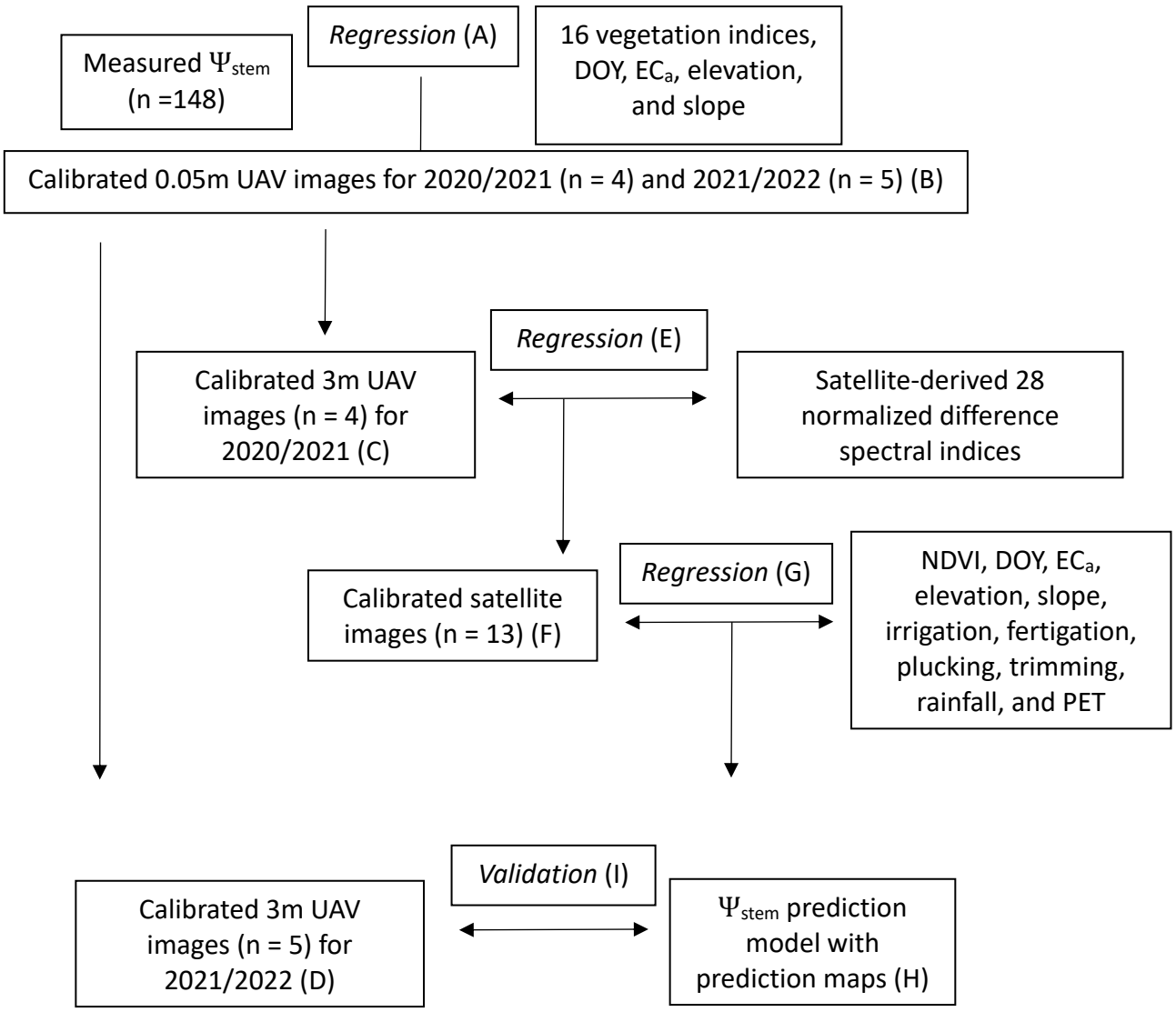


Figure 5. 3 Overview of the workflow. UAV is uncrewed aerial vehicle; DOY is day of the year; NDVI is normalized difference vegetation index; $EC_a$ is apparent electrical conductivity; PET is potential evapotranspiration.

5.2.9.1 Calibration of 0.05m UAV Images

Prior to the first calibration, models were developed by regressing the corresponding 16 VIs, DOY, $EC_a$, elevation, and slope against measured $\Psi_{stem}$ data using RFR, SVR, and MLP (Figure 5.3 (A)). For each grapevine trunk position, the mean values of $EC_a$, elevation, slope, and 16 vegetation indices based on the pure canopy pixels within 0.5 m distance of the trunk were computed for modeling. The acquisition of specific canopy pixels was carried out by overlapping the buffer zones (using recorded trunk location as the center of a circle with a radius of 0.5 m) with the binary raster of canopy height described in section 5.2.3.

The models with the best performance (in terms of RMSE on the test sets) were developed separately for the first growing season (2020/2021) and the second growing season (2021/2022). These models were then used to calibrate the UAV images acquired during 2020/2021 (n = 4) and 2021/2022 (n = 5) (Figure 5.4 (A)), respectively, to obtain calibrated UAV images exhibiting $\Psi_{stem}$ estimation across the vineyards at 0.05m resolution (Figure 5.3 (B); Figure 5.4 (B)).

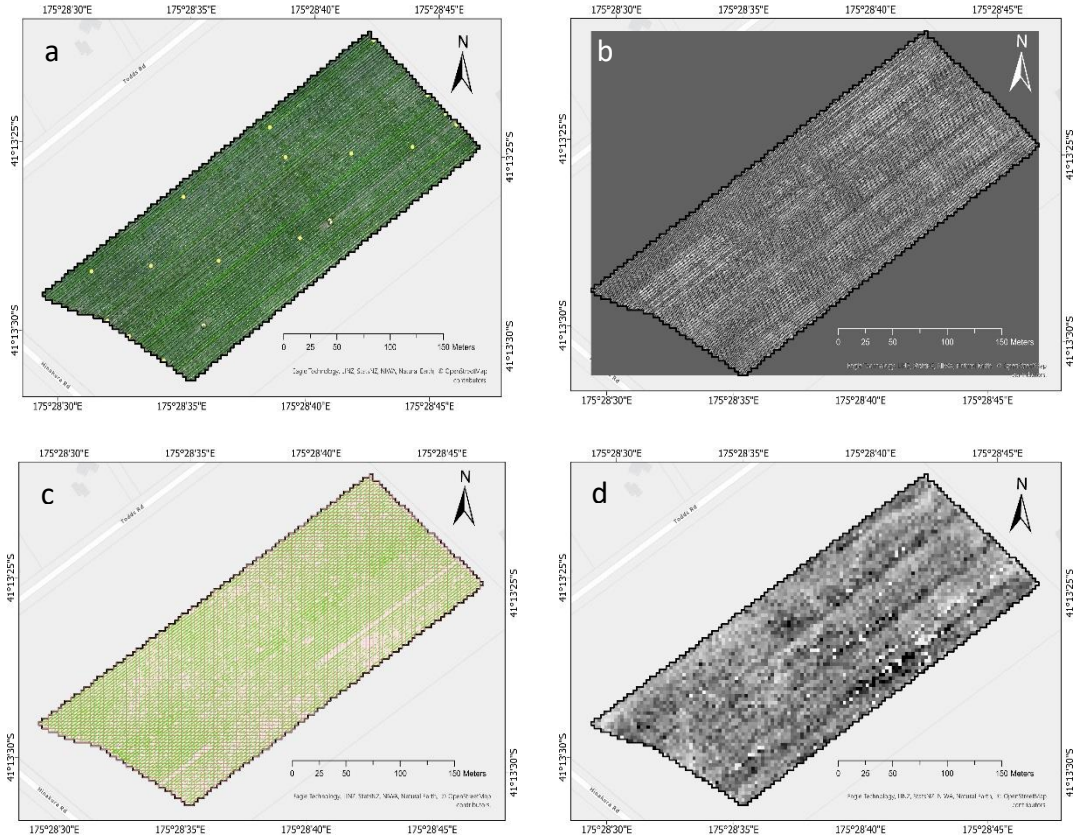



Figure 5. 4 An illustration of generating a downscaled 3m UAV image for each UAV survey date. 4a: the UAV image is presented as RGB composites, with measured $\Psi_{stem}$ presented as yellow points. 4b: the calibrated 0.05m UAV image exhibiting $\Psi_{stem}$ estimation. 4c: the green pixels are pure pixels representing grapevines after clipping pixels of interrow, overlaid with 3m red grids. 4d: the downscaled 3m UAV image after spatial aggregation of pixels in red grids.

#### 5.2.9.2 Generation of Downscaled 3m UAV Images

To proceed with the second calibration for UAV-satellite image pairs (Table 5.1), image co-registration is an essential step. Since UAV images are more spatially accurate due to correction by ground control points, it is a good practice to register the PS images onto the UAV images. In ArcGIS Pro, the UAV images were first resampled from 0.05 m to 3 m resolution and were then used as reference images for the alignment of the PS images. All PS images were then re-aligned with the resampled 3m UAV images using "Snap Raster". Grids of the 3 m pixels were generated by vectorizing the reference images using "Raster to Point" and "Create Thiessen Polygons". As the pixel size of the PS images (3 m) was an integer multiple of that of the UAV images (0.05 m), every PS pixel corresponded to 3,600 UAV pixels after co-registration.

To remove the pixels that did not represent grapevines, the canopy polygons were created based on the binary raster of canopy height obtained in section 2.3 using "Raster to Polygon" for each UAV survey day. These canopy polygons were used to clip the corresponding calibrated UAV images obtained in section 5.2.9.1 to acquire canopy pixels representing grapevines (Figure 5.4 (C)). Those UAV canopy pixels residing in the same cell of the 3m grids were averaged using "Spatial Join". Nine downscaled 3m UAV images (four images in 2020/2021 and five images in 2021/2022) were created (Figure 5.3 (C and D); Figure 5.4 (D)). These downscaled UAV images, representing $\Psi_{stem}$ variability across the vineyards at 3 m resolution, served as references for the calibration of PS images in 2020/2021 (Figure 5.3 (E)) and for the validation of $\Psi_{stem}$ prediction in 2021/2022 (Figure 5.3 (I)).

#### 5.2.9.3 Calibration of satellite images

For the second calibration, models were developed by regressing 28 normalized difference spectral indices (NDSIs) derived from every pixel of the PS images (n = 4; indicated by orange outlines in Table 5.1) against the reference $\Psi_{stem}$ values derived from the 3m calibrated UAV images (n = 4) acquired in 2020/2021 using RFR, SVR, and MLP (Figure 5.3 (E)). NDSIs were computed using all possible pairwise-band combinations in the PS images using "combinations" from the itertools package in Python. The UAV-satellite image pairs had at most one day apart in acquisition date to minimize spectral differences (Table 5.1). The developed models with the best performance (in terms of RMSE on the test sets) were used to calibrate all the PS images (n = 13) acquired during 2020/2021. The series of calibrated satellite images (Figure 5.3 (F)), representing $\Psi_{stem}$ estimation at 3 m resolution, was subsequently used as the response variable in the development of the $\Psi_{stem}$ prediction model.

#### 5.2.9.4 Development of $\Psi_{stem}$ Prediction Model

$\Psi_{stem}$ prediction models were developed by regressing DOY, NDVI, $EC_a$, elevation, slope, total daily rainfall, total daily PET, irrigation, fertigation, plucking, and trimming events, against the reference $\Psi_{stem}$ values derived from the calibrated satellite images (Figure 5.3 (G)). NDVI was selected because it is a widely used proxy for vegetation performance which can partially explain the variation in GWS (Willwerth & Reynolds, 2020). The underlying assumption of the $\Psi_{stem}$ prediction model is that $EC_a$, elevation, and slope are static during the growing seasons by taking advantage of the temporal stability in the spatial patterns of both NDVI and $EC_a$ (Heil & Schmidhalter, 2017; Kazmierski et al., 2011). NDVI needs to be collected at the beginning of each growing season. During any period in the growing season, managers then rely on weather

forecasts for estimating total daily irrigation and PET as well as expected cultivation practices (including irrigation, fertigation, plucking, and trimming) to predict the variation of $\Psi_{stem}$ in vineyards.

NDVI, $EC_a$, elevation, and slope were computed based on the pure canopy pixels within each cell of the 3 m grids. Total daily rainfall, total daily PET, irrigation, fertigation, plucking, and trimming on each day of the 30-day period before the date of $\Psi_{stem}$ prediction were used as independent predictors. Plucking and trimming events were noted using one-hot encoding (1 to mark the event and 0 if the event did not occur). The total number of predictor variables was 125 (Table 5.3).

Table 5. 3 A summary for the predictors used in developing the $\Psi_{stem}$ prediction model. DOY is day of the year; NDVI is normalized difference vegetation index; $EC_a$ is apparent electrical conductivity; PET is potential evapotranspiration.

| Predictor | Note | Number |
|---|---|---|
| DOY | The number is added on along the growing season. | 1 |
| NDVI | Collected in late November for the 2020/2021 and 2021/2022 seasons separately. | 1 |
| $EC_a$ | - | 1 |
| Elevation | - | 1 |
| Slope | - | 1 |
| Total daily rainfall | The rainfall amounts on each day of the 30-day period beforehand was used as a predictor. | 30 |
| Total daily PET | The PET amounts on each day of the 30-day period beforehand was used as a predictor. | 30 |
| Irrigation and Fertigation | The water input amounts, either sourced from irrigation or fertigation, on each day of the 30-day period beforehand was used as a predictor. | 30 |
| Plucking and Trimming | The occurrence of the events on each day of the 30-day period beforehand was used as a predictor. | 30 |

## 5.3 Results

### *5.3.1 Results Variation of Measured $\Psi_{stem}$ Values*

Both vineyards were visited nine times over two growing seasons, from flowering in late November to veraison in late January. Figure 5.5 displays the variability in $\Psi_{stem}$ collected from 148 canopies at the two vineyards, and the distribution of the measurements on each date. The maximum and minimum observation of $\Psi_{stem}$ is 1344 and 310 kPa in 2020/2021, and 1086 and 293 kPa in 2021/2022, respectively. Overall, there is an increasing trend of dehydration in GWS with time for both growing seasons, indicating the increasing presence of water deficit in the grapevines. The only exception is an increase in hydration on 9$^{th}$ December 2021 (compared to the previous measurement) due to rainfall events before sampling (i.e., 38.8 mm on 6th December 2021 and 11 mm on 7th December 2021). In Figure 5.5, the height of the box (the difference between the upper quartile and lower quartile) represents the spatial

variation of $\Psi_{stem}$ on one date across the relevant vineyard. This difference increased as the survey proceeded, implying spatial variation within the vineyards became more pronounced as water deficit increased. The mean values of $\Psi_{stem}$ obtained at Pencarrow were lower than those obtained at Wharekauhau on most of the measurement dates.

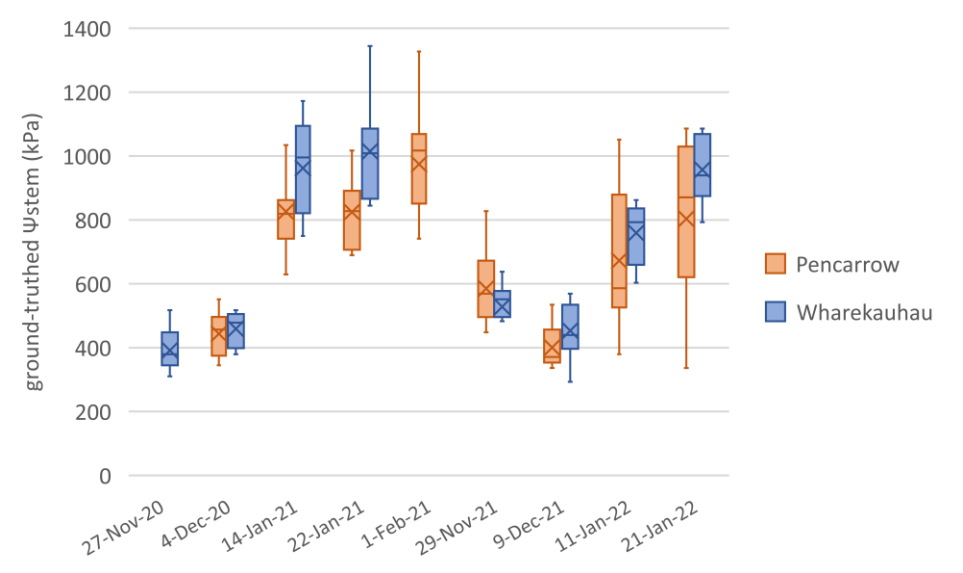



Figure 5. 5 Box plot of measured $\Psi_{stem}$ collected during two growing seasons at Pencarrow (n = 86) and Wharekauhau (n = 62) vineyards. X symbols refer to the average values on the survey dates. Horizontal lines in the boxes refer to median values on the survey dates.

*5.3.2 Modeling Performance*

The regression modeling was conducted for the two-stage calibration (Figure 5.3 (A and E)) and the development of the $\Psi_{stem}$ prediction model (Figure 5.3 (G)) using RFR, SVR, and MLP. The ML models with the best performance (in terms of RMSE on the test sets) were chosen (Table 5.4). The RMSE of the test sets for two-stage calibration, ranging from 59 to 113 kPa, indicates that the estimation obtained from the calibration models is well correlated with measured $\Psi_{stem}$ (first calibration) and reference $\Psi_{stem}$ (second calibration; Figure 5.3 (C)). Therefore, the calibrated satellite images (Figure 5.3 (F)) should serve as a robust reference as response variables for $\Psi_{stem}$ prediction (Figure 5.3 (G)). The RMSE of the test sets for $\Psi_{stem}$ prediction is 31 kPa, suggesting the prediction model adequately captured $\Psi_{stem}$ variability in 2020/2021. The scatter plots (Figure 5.6) show that the data distribution for each regression model is closely clustered around the 1:1 line.

Table 5. 4 Results of modeling performance. UAV is uncrewed aerial vehicle; $\Psi_{stem}$ is stem water potential; RMSE is root mean square error. The data size for "satellite images" and "prediction of $\Psi_{stem}$" indicates the number of pixels used in the regression modeling. RPIQ is ratio of performance to interquartile range.

| Modeling Purpose | Regression Model | Data Size of Measured or Reference $\Psi_{stem}$ | RMSE of the Training Set (kPa) | RPIQ of the Training Set (kPa) | RMSE of the Test Set (kPa) | RPIQ of the Test Set (kPa) |
|---|---|---|---|---|---|---|
| Calibration of UAV images acquired in 2020/2021 (Figure 5.3 (A)) | Multilayer perceptron | 85 | 96 | 5.5 | 113 | 4.2 |
| Calibration of UAV images acquired in 2021/2022 (Figure 5.3 (A)) | Random forest regression | 63 | 121 | 3.5 | 106 | 2.2 |
| Calibration of Satellite images (Figure 5.3 (E)) | Random forest regression | 42,234 | 47 | 5.5 | 59 | 4.4 |
| Prediction of $\Psi_{stem}$ (Figure 5.3 (G)) | Random forest regression | 151,580 | 25 | 9.2 | 31 | 7.5 |

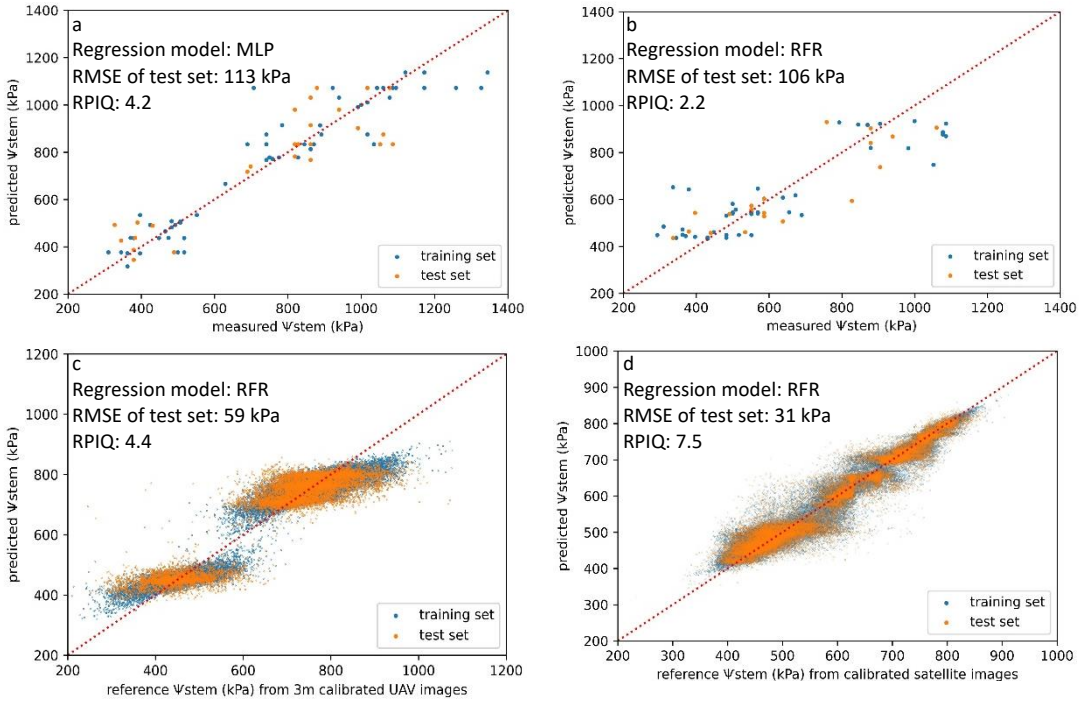



Figure 5. 6 Scatter plots between predicted $\Psi_{stem}$ and measured or reference $\Psi_{stem}$ (kPa) for the training and test sets of 6a: the calibration of UAV images acquired in 2020/2021. 6b: the calibration of UAV images acquired in 2021/2022. 6c: the calibration of satellite (PS) images. 6d: the prediction of $\Psi_{stem}$ in 2020/2021. Red dashed lines are 1:1 lines. Samples from the two study vineyards are considered collectively in each regression model.

### *5.3.3 The Accuracy of the Spatial and Temporal Patterns of $\Psi_{stem}$ Prediction*

An important goal of this study was to validate whether a $\Psi_{stem}$ prediction model based only on the first-season data (2020/2021) can be used to predict $\Psi_{stem}$ changes in the second season (2021/2022). The similarity analysis, between $\Psi_{stem}$ prediction maps (obtained from the best prediction model) and the $\Psi_{stem}$ reference maps (obtained from 3m calibrated UAV images in 2021/2022) was conducted using Pearson correlation for each vineyard. The values of r represent the degree of consistency between predicted and reference maps across multiple survey dates, with p-values showing a significant association (Table 5.5).

Figure 5.7 shows $\Psi_{stem}$ prediction alongside measured $\Psi_{stem}$ values (collected by grid sampling throughout the study periods in 2021/2022) during the season for both vineyards. $\Psi_{stem}$ predictions were normalized for better visualization. Both $\Psi_{stem}$ prediction series for both vineyards appear to intercept with each of the spreads of measured $\Psi_{stem}$ data approximately at their average value, except for the last measurements at Wharekauhau vineyard. Note that higher $\Psi_{stem}$ values indicate more dehydrated vines. Measured $\Psi_{stem}$ values are assumed to represent the total variability of each vineyard. Thus, this result indicates that the temporal patterns in 2020/2021 captured by the $\Psi_{stem}$ prediction model can provide good temporal predictions during the study period in 2021/2022.

Table 5. 5 Results of similarity analysis presented by the Pearson correlation coefficient (r), represent the consistency between predicted and reference $\Psi_{stem}$ maps across four survey dates for the two study vineyards.

| | Pencarrow | Wharekauhau |
|---|---|---|
| r | 0.89* | 0.87* |

Significance levels are noted as * when $p \leq 0.001$

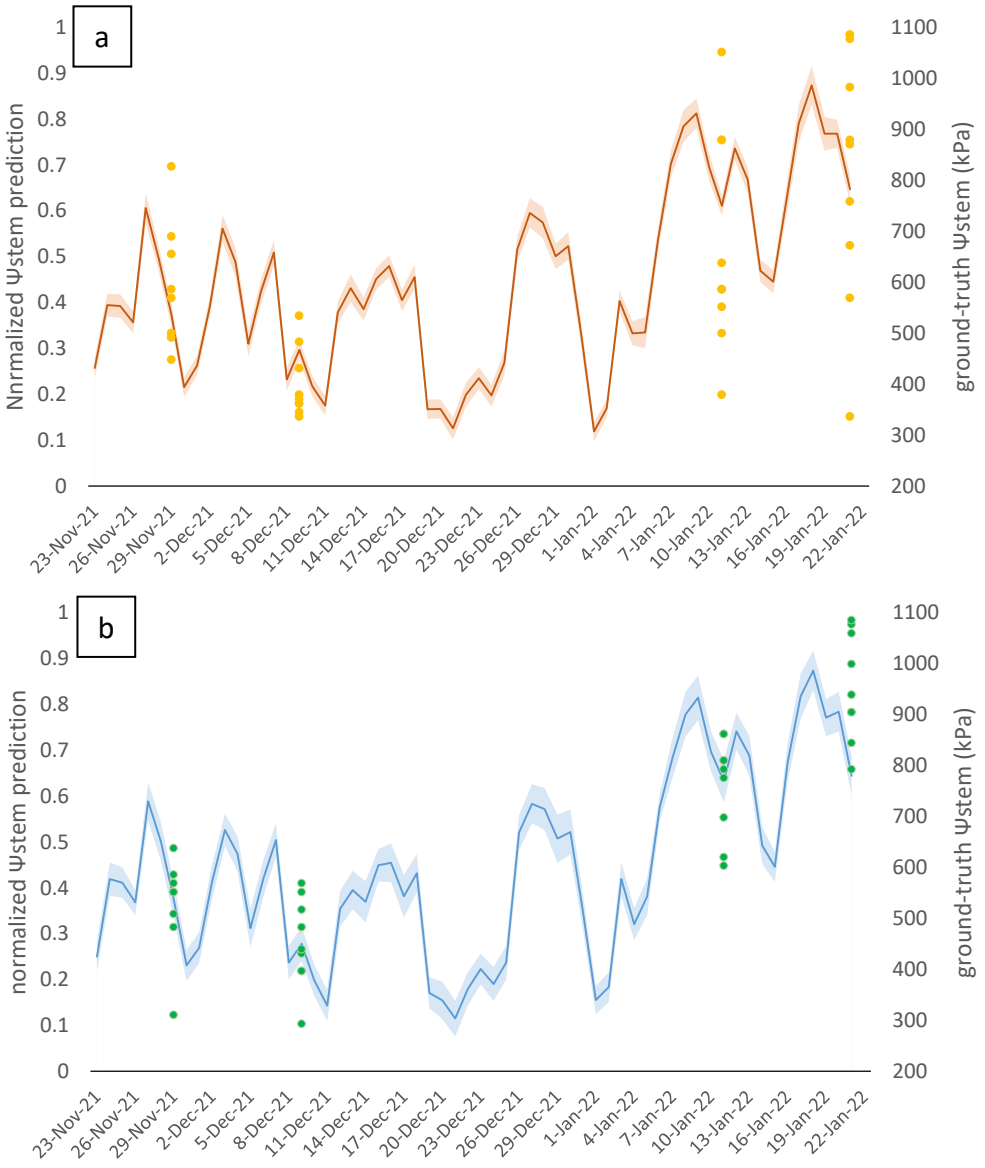



Figure 5. 7 The temporal patterns of normalized $\Psi_{stem}$ prediction (lines) during the growing season in 2021/2022, with the measured $\Psi_{stem}$ measurements (points) for Pencarrow (5.7a) and Wharekauhau (5.7b). The shaded bands bordering the lines encompass one standard deviation.

Descriptive statistics, including mean, standard deviation (SD), and coefficient of variation (CV), are presented in Table 5.6 to help assess the 2021/2022 datasets of measured $\Psi_{stem}$ values, reference $\Psi_{stem}$ values obtained from 3m calibrated UAV images, and predicted $\Psi_{stem}$ values generated by the $\Psi_{stem}$ prediction model. The statistics indicate that measured $\Psi_{stem}$ values (CV > 10%) were more heterogeneous than the $\Psi_{stem}$ values in the corresponding 3m calibrated UAV datasets (4% > CV > 2%) and the prediction datasets (CV < 2%). Thus, the tendency to exhibit extreme $\Psi_{stem}$ values decreases from the 3m calibrated UAV datasets (n = 42234, $\Psi_{stem\ max}$ = 867, $\Psi_{stem\ min}$ = 473) to the prediction datasets (n = 151580, $\Psi_{stem\ max}$ = 694, $\Psi_{stem\ min}$ = 550) compared to the measured datasets (n = 63, $\Psi_{stem\ max}$ = 1086 kPa, $\Psi_{stem\ min}$ = 293 kPa). Although there are differences in the absolute values and variability of $\Psi_{stem}$ for the three datasets, they approximately follow the same temporal patterns over the study

period in 2021/2022 (Figure 5.8). This indicates that both 3m calibrated UAV datasets and $\Psi_{stem}$ prediction datasets are more reliable predictors of the temporal patterns of $\Psi_{stem}$, rather than their absolute values.

Table 5. 6 Summary statistics for predicted $\Psi_{stem}$, $\Psi_{stem}$ acquired from 3m calibrated UAV images, and measured $\Psi_{stem}$ for each survey date at the Pencarrow and Wharekauhau vineyards. UAV is uncrewed aerial vehicle; SD is standard deviation; CV is coefficient of variation.

| Survey Date | Vineyard | Data Source | Mean | SD | CV |
|---|---|---|---|---|---|
| 29th November 2021 | Pencarrow | Predicted $\Psi_{stem}$ | 609 | 5.59 | 0.92 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 562 | 18.11 | 3.22 |
| | | Measured $\Psi_{stem}$ | 585 | 113.93 | 19.47 |
| | Wharekauhau | Predicted $\Psi_{stem}$ | 603 | 9.26 | 1.53 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 565 | 21.86 | 3.87 |
| | | Measured $\Psi_{stem}$ | 528 | 87.31 | 16.55 |
| 09th December 2021 | Pencarrow | Predicted $\Psi_{stem}$ | 593 | 5.03 | 0.85 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 547 | 15.19 | 2.78 |
| | | Measured $\Psi_{stem}$ | 400 | 64.25 | 16.05 |
| | Wharekauhau | Predicted $\Psi_{stem}$ | 580 | 8.95 | 1.54 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 548 | 17.42 | 3.18 |
| | | Measured $\Psi_{stem}$ | 453 | 82.18 | 18.14 |
| 11th January 2022 | Pencarrow | Predicted $\Psi_{stem}$ | 662 | 4.84 | 0.73 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 681 | 19.53 | 2.87 |
| | | Measured $\Psi_{stem}$ | 672 | 204.27 | 30.39 |
| 21st January 2021 | Pencarrow | Predicted $\Psi_{stem}$ | 670 | 6.73 | 1.01 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 717 | 22.86 | 3.19 |
| | | Measured $\Psi_{stem}$ | 803 | 233.27 | 29.03 |
| | Wharekauhau | Predicted $\Psi_{stem}$ | 668 | 10.35 | 1.55 |
| | | $\Psi_{stem}$ from 3m calibrated UAV | 740 | 24.67 | 3.34 |
| | | Measured $\Psi_{stem}$ | 957 | 99.35 | 10.39 |

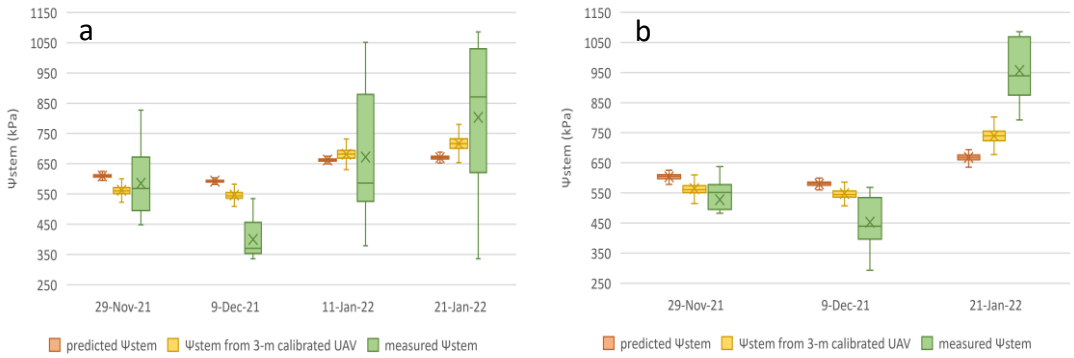



Figure 5. 8 Box plots of predicted $\Psi_{stem}$ values, $\Psi_{stem}$ values acquired from 3m calibrated UAV images, and measured $\Psi_{stem}$ values for each survey date at the Pencarrow (8a) and Wharekauhau (8b) vineyards. UAV is uncrewed aerial vehicle, and $\Psi_{stem}$ is stem water potential.

## 5.4 Discussion

### *5.4.1 Reference $\Psi_{stem}$ Data Generated by Two-Stage Calibration*

To generate reference $\Psi_{stem}$ data for subsequent prediction modeling, two-stage calibration was employed to calculate $\Psi_{stem}$ data from satellite images based on measured $\Psi_{stem}$ data using RFR, SVR, and MLP. Satellite images (obtained from PS in this study) have advantages over UAV data or measured data, in terms of higher coverage during one acquisition and regular acquisition periods. Therefore, satellite images are a better source in terms of providing observations on a regular basis. However, when assessing vertically-oriented grapevine status with discontinuous layouts, the utilization of low-to-moderate satellite resolution is often challenging due to the biased estimation resulting from pixels with mixed signals (Borgogno-Mondino, Lessio, et al., 2018). To address the issue of mixed-signal pixels, two calibration events were applied to scale the measured measurements up to data at the 3 m satellite level. UAV-based data served a critical role in this approach as it enhanced spectral purity by removing non-grapevine pixels and provided a large number of 0.05m reference data, which was then spatially aggregated to ensure scale matching between satellite data and measured measurements (Pla et al., 2019; Revill et al., 2020). The performance (in terms of RMSE on the test sets) of the first and the second calibration for the data acquired in 2020/2021 was 113 and 59 kPa, respectively (Table 5.4). These results supported the appropriateness of this two-stage calibration approach in generating large reference datasets (n = 151580), compared to the measured samples (n = 85), for later modeling purposes.

### *5.4.2 Predicted $\Psi_{stem}$ Data Generated by $\Psi_{stem}$ Prediction Model*

The $\Psi_{stem}$ prediction model was developed using calibrated PS images acquired in 2020/2021 as response variables, and using NDVI, $EC_a$, elevation, slope, DOY, irrigation, fertigation, plucking, and trimming events, daily total rainfall, and daily total PET as predictors. The need for this $\Psi_{stem}$ prediction model was demonstrated by the number of PS acquisition images in this study. As the revisit period of PS is daily, there should potentially be 80 images, with surface reflectance assets, proper ground control, and standard quality, available over the study periods during 2020/2021 (15$^{th}$ November 2020 to 2$^{nd}$ February 2021). Nevertheless, there were only 20 images available to download, with only 13 images (the longest time gap between acquisition

dates being 17 days) suitable for analysis (after eliminating those images contaminated with cloud, haze, or cloud shadow). Bellvert et al. (2016) have suggested access to five to six points of information acquisition over the season for vineyard irrigation scheduling when considering cost efficiency. To avoid inaccessible satellite data caused by weather or technical problems, the development of a GWS prediction model is desirable.

The trained model performed well on the test set of data obtained in 2020/2021, with an RMSE of 31 kPa on the test sets (Table 5.4). From the two-stage calibration to $\Psi_{stem}$ prediction, it appears the accuracy of modeling was higher as the data size of measured or reference $\Psi_{stem}$ increased, resulting from the two-stage calibration. This confirmed that ML models develop more robust relationships with fewer overfitting issues when they are trained with more samples (Lever et al., 2016). The superior performance of RFR with calibration and prediction modeling is probably because RFR is robust to multicollinearity (Zhan et al., 2018). Multicollinearity was evident in the datasets of this study, originating from VI computation based on the same spectral band, and temporal autocorrelation between the rainfall and PET data series (Appendix 9).

In this study, the model was further evaluated with the data independently collected during a second growing season (2021/2022) to test its generalization performance. The model reflected the trends of changes of $\Psi_{stem}$ in response to measured $\Psi_{stem}$. The predicted temporal patterns became more hydrated in early December compared to late November but exhibited increasing water deficit as the growing season proceeded, which is comparable to measured $\Psi_{stem}$ data (Figures 5.5 and 5.7). Similarity analysis (in terms of r) indicated that temporal patterns were well depicted by the prediction models for both vineyards (Table 5.5). It should be noted that the model's prediction capability was evaluated using calibrated UAV images (reference maps) instead of using measured $\Psi_{stem}$. The reasons of that include the outputs of the prediction model are of 3-m spatial resolution, and each pixel corresponds to multiple grapevines. Thus, the outputs cannot be directly compared with individually measured $\Psi_{stem}$. Besides, PS images acquired in 2021/2022 cannot be used for evaluation because background information (soil and grass) is included. Therefore, the prediction model was evaluated using the calibrated UAV images that approximated the measured $\Psi_{stem}$ assumed to represent the total GWS variability across the study vineyards (RMSE = 106 kPa). The results of similarity analysis be attributed to the observed similarity in PET variability over the two seasons (Figure 5.2). As ET is a critical component in the soil water balance contributing to $\Psi_{stem}$ (Montoro et al., 2006), the similarity of ET from different seasons may lead to a similarity in the temporal variation of $\Psi_{stem}$ over different years. Another reason for the similarity is that cultivation practices were applied within similar timeframes during each growing season, so their impacts were still within the prediction range of the model. In addition, both NDVI-defined and $EC_a$-defined zones have been reported to support spatial assessment of GWS (Acevedo-Opazo et al., 2008; Yu et al., 2020), further increasing the prediction capabilities of the model.

The prediction model established the temporal variation of $\Psi_{stem}$ based on DOY, 30-day total daily rainfall, 30-day total daily PET, 30-day irrigation, 30-day fertigation, 30-day plucking, and 30-day trimming, while accounting for the spatial variation of $\Psi_{stem}$ based on NDVI, $EC_a$, elevation, and slope. Although temporal patterns of $\Psi_{stem}$ were tested with promising results, it should be noted that the relative pattern of values, rather than their absolute values, should be the focus of attention in this study. Therefore, sampling for Ψstem prediction should be undertaken at the beginning of

each growing season for model calibration, so as to enhance the precision and practicability of the model for use in irrigation scheduling.

### *5.4.3 Limitations and Directions for Future Research*

The weakness of the prediction model in this study lies in its empirical approach, so it is unable to provide a reliable prediction if the inputs are beyond the range of observed variability. Field measurements are needed to validate the prediction model with data collected from additional phenological stages, growing seasons, and sites, to enhance the model's scope. In addition, due to the selection of 3m resolution satellite images, high heterogeneity of GWS within a cell of grids would be averaged and exhibit less variation in $\Psi_{stem}$, thus introducing bias into calibration and modeling.

In its current form, this prediction model will only be able to predict $\Psi_{stem}$ at vineyard-block scale during the growing season, rather than provide a spatial map of $\Psi_{stem}$ each day. One reason is that weather predictors (including daily total rainfall, and daily total PET) were assumed to be evenly spatially distributed across the two vineyards. Predictors accounting for spatial variation (including NDVI, $EC_a$, elevation, and slope) are static during modeling and are determined before the growing seasons. Despite these limitations, this prediction model is still a potentially practical tool for viticulturists since the block size of the results can be adjusted, by increasing the number of weather sensors or stations, to suit the size of zones that need different irrigation management.

Subsequent research is suggested to explore new predictor variables that are cost-effective to acquire and can serve as spatio-temporal proxies for $\Psi_{stem}$. Weather parameters, recorded spatially and continuously by wireless sensor networks across vineyards, could be one of these potential predictors. Further research work should examine the number and type of field measurements required and protocols needed for vineyard sampling to contribute to model calibration at the beginning of the growing season. The impact of the quality of weather forecasts on the performance of the prediction model also requires investigation. This will, potentially, allow the model to become a more reliable prediction tool, providing daily $\Psi_{stem}$ spatial maps in advance throughout the growing season.

## 5.5 Conclusions

This study demonstrated the potential application of using a two-stage calibration approach for calibrating satellite images to provide reference stem water potential ($\Psi_{stem}$) data. The potential of establishing a $\Psi_{stem}$ prediction model based on day of the year, normalized difference vegetation indices, apparent electrical conductivity, elevation, slope, rainfall, potential evapotranspiration, irrigation, fertigation, plucking, and trimming events, was also demonstrated. Collection of ground truthing is required for model calibration at the beginning of the growing season. The prediction model can be improved on a daily basis if predictors that account for spatio-temporal variability in $\Psi_{stem}$ are provided, such as spatially recorded weather information. This tool has the potential to benefit vineyard managers with improved irrigation management and quality optimization by providing $\Psi_{stem}$ predictions at vineyard-block scale during growing seasons, when the model is properly calibrated and coupled with accurate weather forecasts.

**References**


Acevedo-Opazo, C., Tisseyre, B., Guillaume, S., & Ojeda, H. (2008). The potential of high spatial resolution information to define within-vineyard zones related to vine water status. *Precision Agriculture, 9*(5), 285-302. https://doi.org/10.1007/s11119-008-9073-1

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome, 300*(9), D05109.

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*.

Baluja, J., Diago, M. P., Balda, P., Zorer, R., Meggio, F., Morales, F., & Tardaguila, J. (2012). Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV). *Irrigation Science, 30*(6), 511-522. https://doi.org/10.1007/s00271-012-0382-9

Barnes, E., Clarke, T., Richards, S., Colaizzi, P., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., & Thompson, T. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, USA,

Bellvert, J., Marsal, J., Mata, M., & Girona, J. (2012). Identifying irrigation zones across a 7.5-ha 'Pinot noir'vineyard based on the variability of vine water status and multispectral images. *Irrigation Science, 30*(6), 499-509.

Bellvert, J., Zarco-Tejada, P. J., Marsal, J., Girona, J., González-Dugo, V., & Fereres, E. (2016). Vineyard irrigation scheduling based on airborne thermal imagery and water potential thresholds. *Australian Journal of Grape and Wine Research, 22*(2), 307-315.

Birth, G. S., & McVey, G. R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer 1. *Agronomy Journal, 60*(6), 640-643.

Borgogno-Mondino, E., Lessio, A., Tarricone, L., Novello, V., & De Palma, L. (2018). A comparison between multispectral aerial and satellite imagery in precision viticulture. *Precision Agriculture, 19*(2), 195-217.

Borgogno-Mondino, E., Novello, V., Lessio, A., & de Palma, L. (2018). Describing the spatio-temporal variability of vines and soil by satellite-based spectral indices: A case study in Apulia (South Italy). *International Journal of Applied Earth Observation and Geoinformation, 68*, 42-50.

Bramley, R., Ouzman, J., & Boss, P. K. (2011). Variation in vine vigour, grape yield and vineyard soils and topography as indicators of variation in the chemical composition of grapes, wine and wine sensory attributes. *Australian Journal of Grape and Wine Research, 17*(2), 217-229.

Brillante, L., Martínez-Luscher, J., Yu, R., Plank, C. M., Sanchez, L., Bates, T. L., Brenneman, C., Oberholster, A., & Kurtural, S. K. (2017). Assessing spatial variability of grape skin flavonoids at the vineyard scale based on plant water status mapping. *Journal of Agricultural and Food Chemistry, 65*(26), 5255-5265.

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.

Bukowiecki, J., Rose, T., & Kage, H. (2021). Sentinel-2 Data for Precision Agriculture?—A UAV-Based Assessment. *Sensors, 21*(8), 2861.

Chaves, M., Zarrouk, O., Francisco, R., Costa, J., Santos, T., Regalado, A., Rodrigues, M., & Lopes, C. (2010). Grapevine under deficit irrigation: hints from physiological and molecular data. *Annals of Botany, 105*(5), 661-676.

Cook, P. G., & Williams, B. G. (1998). *Electromagnetic Induction Techniques-Part 8*. CSIRO PUBLISHING.

Daughtry, C. S., Walthall, C., Kim, M., De Colstoun, E. B., & McMurtrey Iii, J. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment, 74*(2), 229-239.

Devaux, N., Crestey, T., Leroux, C., & Tisseyre, B. (2019). Potential of Sentinel-2 satellite images to monitor vine fields grown at a territorial scale. *OENO One, 53*(1), 52-59.

Espinoza, C. Z., Khot, L. R., Sankaran, S., & Jacoby, P. W. (2017). High resolution multispectral and thermal remote sensing-based water stress assessment in subsurface irrigated grapevines. *Remote Sensing, 9*(9), 961.

Fernandes-Silva, A., Oliveira, M., Paço, T. A., & Ferreira, I. (2019). Deficit irrigation in Mediterranean fruit trees and grapevines: Water stress indicators and crop responses. *Irrigation in Agroecosystems. London, United Kingdom*.

Gamon, J., & Surfus, J. (1999). Assessing leaf pigment content and activity with a reflectometer. *The New Phytologist, 143*(1), 105-117.

Gautam, D., & Pagay, V. (2020). A review of current and potential applications of remote sensing to study the water status of horticultural crops. *Agronomy, 10*(1), 140.

Giovos, R., Tassopoulos, D., Kalivas, D., Lougkos, N., & Priovolou, A. (2021). Remote sensing vegetation indices in viticulture: A critical review. *Agriculture, 11*(5), 457.

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment, 80*(1), 76-87.

Gitelson, A. A., & Merzlyak, M. N. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research, 22*(5), 689-692.

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters, 32*(8).

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment, 90*(3), 337-352.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment, 81*(2-3), 416-426.

Heil, K., & Schmidhalter, U. (2017). The application of EM38: Determination of soil parameters, selection of soil sampling points and use in agriculture and archaeology. *Sensors, 17*(11), 2540.

Helman, D., Bahat, I., Netzer, Y., Ben-Gal, A., Alchanatis, V., Peeters, A., & Cohen, Y. (2018). Using time series of high-resolution planet satellite images to monitor grapevine stem water potential in commercial vineyards. *Remote Sensing, 10*(10), 1615.

Houborg, R., & McCabe, M. F. (2018). Daily Retrieval of NDVI and LAI at 3 m Resolution via the Fusion of CubeSat, Landsat, and MODIS Data. *Remote Sensing, 10*(6), 890.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*(1-2), 195-213.

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.

Kazmierski, M., Glémas, P., Rousseau, J., & Tisseyre, B. (2011). Temporal stability of within-field patterns of NDVI in non irrigated Mediterranean vineyards. *OENO One, 45*(2), 61-73.

Khaliq, A., Comba, L., Biglia, A., Ricauda Aimonino, D., Chiaberge, M., & Gay, P. (2019). Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. *Remote Sensing, 11*(4), 436.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer.

Leach, N., Coops, N. C., & Obrknezev, N. (2019). Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies. *Computers and Electronics in Agriculture, 164*, 104893.

Lever, J., Krzywinski, M., & Altman, N. (2016). Points of significance: model selection and overfitting. *Nature methods, 13*(9), 703-705.

Martínez-Lüscher, J., Sánchez-Díaz, M., Delrot, S., Aguirreolea, J., Pascual, I., & Gomes, E. (2014). Ultraviolet-B radiation and water deficit interact to alter flavonol and anthocyanin profiles in grapevine berries through transcriptomic regulation. *Plant and Cell Physiology, 55*(11), 1925-1936.

Matese, A., Baraldi, R., Berton, A., Cesaraccio, C., Di Gennaro, S. F., Duce, P., Facini, O., Mameli, M. G., Piga, A., & Zaldei, A. (2018). Estimation of water stress in grapevines using proximal and remote sensing methods. *Remote Sensing, 10*(1), 114.

Matese, A., Toscano, P., Di Gennaro, S., Genesio, L., Vaccari, F., 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*(3), 2971-2990. https://doi.org/10.3390/rs70302971

Mazzia, V., Comba, L., Khaliq, A., Chiaberge, M., & Gay, P. (2020). UAV and machine learning based refinement of a satellite-driven vegetation index for precision agriculture. *Sensors, 20*(9), 2530.

Monaghan, J. M., Daccache, A., Vickers, L. H., Hess, T. M., Weatherhead, E. K., Grove, I. G., & Knox, J. W. (2013). More 'crop per drop': constraints and opportunities for precision irrigation in European agriculture. *Journal of the Science of Food and Agriculture, 93*(5), 977-980.

Montoro, A., López Urrea, R., Mañas, F., López Fuster, P., & Fereres, E. (2006). Evapotranspiration of grapevines measured by a weighing lysimeter in La Mancha, Spain. V International Symposium on Irrigation of Horticultural Crops 792,

Ohana-Levi, N., Munitz, S., Ben-Gal, A., & Netzer, Y. (2020). Evaluation of within-season

grapevine evapotranspiration patterns and drivers using generalized additive models. *Agricultural Water Management, 228*, 105808.

Ojeda, H., Andary, C., Kraeva, E., Carbonneau, A., & Deloire, A. (2002). Influence of pre- and postveraison water deficit on synthesis and concentration of skin phenolic compounds during berry growth of Vitis vinifera cv. Shiraz. *American Journal of Enology and Viticulture, 53*(4), 261-267.

Patakas, A., Noitsakis, B., & Chouzouri, A. (2005). Optimization of irrigation water use in grapevines using the relationship between transpiration and plant water status. *Agriculture, Ecosystems & Environment, 106*(2-3), 253-259.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass.*Series A. Mathematical Physical Sciences* Proceedings of the Royal Society of London,

Pla, M., Bota, G., Duane, A., Balagué, J., Curcó, A., Gutiérrez, R., & Brotons, L. (2019). Calibrating Sentinel-2 imagery with multispectral UAV derived information to quantify damages in Mediterranean rice crops caused by Western Swamphen (Porphyrio porphyrio). *Drones, 3*(2), 45.

Planet Team. (2017). Planet application program interface: In space for life on earth. *San Francisco*.

Poblete, T., Ortega-Farias, S., Moreno, M. A., & Bardeen, M. (2017, Oct 30). Artificial Neural Network to Predict Vine Water Status Spatial Variability Using Multispectral Information Obtained from an Unmanned Aerial Vehicle (UAV). *Sensors (Basel), 17*(11). https://doi.org/10.3390/s17112488

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment, 48*(2), 119-126.

Revill, A., Florence, A., MacArthur, A., Hoad, S., Rees, R., & Williams, M. (2020). Quantifying uncertainty and bridging the scaling gap in the retrieval of leaf area index by coupling Sentinel-2 and UAV observations. *Remote Sensing, 12*(11), 1843.

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. https://doi.org/10.1016/j.compag.2018.02.013

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55*(2), 95-107.

Rouse, J., Haas, R., Schell, J., Deering, D., & Harlan, J. (1974). Monitoring the vernal advancements and retrogradation. *Texas AM Univ, Texas*.

Schreiner, R. P., & Lee, J. (2014). Effects of post-véraison water deficit on 'Pinot noir'yield and nutrient status in leaves, clusters, and musts. *HortScience, 49*(10), 1335-1340.

Sozzi, M., Kayad, A., Marinello, F., Taylor, J., & Tisseyre, B. (2020). Comparing vineyard imagery acquired from Sentinel-2 and Unmanned Aerial Vehicle (UAV) platform. *OENO One, 54*(2), 189-197. Suter, B., Triolo, R., Pernet, D., Dai, Z., & Van Leeuwen, C. (2019). Modelling stem water potential by separating the effects of soil water availability and climatic conditions on water status in grapevine (Vitis vinifera L.). *Frontiers in Plant Science, 10*, 1485.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*(2), 127-150.

van LEEUWEN, C., GOUTOULY, J.-P., COSTA-FERREIRA, A.-M., AZAÏS, C., MARGUERIT, E., ROBY, J.-P., CHONE, X., GERMAIN, C., HOMAYOUNI, S., & GAUDILLERE, J.-P. (2006). Intra-block variations of vine water status in time and space Variations intra-parcellaires temporelles et spatiales du régime hydrique de la vigne. *Vine*, 64-69.

Van Leeuwen, C., Trégoat, O., Choné, X., Bois, B., Pernet, D., & Gaudillère, J.-P. (2009). Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? *OENO One, 43*(3), 121-134.

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*(16), 3198.

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*(23), 5918.

Willwerth, J. J., & Reynolds, A. G. (2020). Spatial variability in Ontario Riesling Vineyards: I. Soil, vine water status and vine performance. *OENO One, 54*, 311-333.

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE, 38*(1), 259-269.

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*.

Zhan, Y., Luo, Y., Deng, X., Zhang, K., Zhang, M., Grieneisen, M. L., & Di, B. (2018). Satellite-based estimates of daily NO2 exposure in China using hybrid random forest and spatiotemporal kriging model. *Environmental science & technology, 52*(7), 4180-4189.

Chapter 5 demonstrates the proof of improving the application issues when employing multispectral satellite images in viticultural water management. The variability contained in ground measurements can be reliably upscaled to satellite images by implementing two-level spatial calibration using UAV data as an interim step. The temporal patterns of block GWS subsequently can be predicted when using environmental variables (including soil/terrain data, weather parameters, temporal information, and cultivation practices) as predictors. Considering these chapters as a whole, the provision of vineyard-scale or regional-scale GWS prediction is established based on the following findings: identification of GWS-sensitive spectral bands, determination of ancillary effects provided by environmental factors, proof of the usefulness of two-stage calibration for upscaling ground information, and development of the prediction model according to the weather forecast.

# Chapter 6
# General Discussion

## 6.1 A Summary of Main Findings from Each Paper

Wine grapes are an important horticultural crop in the world in terms of the planted area, production volume, consumption, and export value (Chapter 1). In NZ, wine is the second-highest horticultural export by value in 2021, after kiwifruit. Nevertheless, according to the four developed scenarios in the IPCC report (IPCC, 2014), ongoing climate change poses foreseeable challenges to agricultural activities. NZ will experience higher mean temperatures, higher levels of radiation, fluctuation in rainfall patterns, and more occurrences of drought and heatwaves in the projected long-term trends (Solomon et al., 2007; Stocker, 2014). These will lead to uncertainty in grape quality associated with the variation in vegetation water status, as moderate water deficit in a narrow range is beneficial to grape composition and wine properties (-1200 or -1300 kPa for red varieties after fruit set) (Baeza et al., 2019; Fernandes-Silva et al., 2019). Meanwhile, the amended Resource Management Regulations increased the frequency of water take reporting and minimum flow rates from rivers, with capped water allocation in NZ. This further exacerbates the competition for water requirements among NZ industrial sectors. Given the upcoming climate changes, stricter water take regulations, and the importance of adjusting dehydration, there is a desire for the viticultural industry to accurately estimate the level of water stress before applying irrigation to achieve optimal grape quality.

Recent advances in the availability of remotely sensed information, geospatial techniques, and statistical modeling have enabled the estimation of water deficit across vineyards and along growing seasons. However, there are challenges to the application and adoption of these technologies, and limited studies undertaken to address these knowledge gaps. This study investigated the RS and geospatial options for precision irrigation, specifically for improving regulated DI practices in vertical-oriented Pinot Noir vineyards in Martinborough, New Zealand. Technical limitations and bottlenecks related to the use of spectral information in GWS estimation were highlighted and tackled from three aspects: (i) spectral data analysis and processing pipelines, (ii) the interoperability among multi-sourced data, and (iii) the satellite imagery with high temporal and spatial resolution. The main part of the discussion is governed by these three aspects which are composed of research objectives, a brief introduction, applied methods, and summarized findings. New information and novel solutions are provided to address the issues met, when applying proximally sensed hyperspectral data and remotely sensed multispectral images to GWS estimation.

### *6.1.1 The Relationships between $\Psi_{stem}$ and Spectral Data*

The goal of using remotely sensed spectral information as a proxy for in situ measurements of GWS is to provide accurate and large-scale estimations to assist in timely or in advance irrigation scheduling. In this study, $\Psi_{stem}$ is selected as the reliable ground truth for GWS substitute, as it considers the integrated effects of environmental drivers on vegetation water status while being a sensitive indicator to sense early dehydration in the grapevines (Choné, 2001). To this end, the relationships

between spectral data and $\Psi_{stem}$ are required for a robust estimation of GWS. This study was designed to detect moderate water deficit since the $\Psi_{stem}$ variance of 97.9 % of the samples spans across non-stress ($\Psi_{stem} \leq 800$ kPa) to moderate stress (800 kPa $\leq \Psi_{stem} \leq 1200$ kPa) (Acevedo-Opazo et al., 2008; Appendix 1). Systematic sampling was carried out to select sampled canopies (Appendix 10). The locations of several posts were selected and recorded as the ground control points to increase the spatial accuracy of UAV and satellite images (Appendix 10). Descriptive statistics for training and test set used in Chapter 3, 4, and 5 are presented in Appendix 5. It is noted that although the values of mean and standard deviation between the training and test set in Chapter 3, 4, and 5 are close, the value range of the training set does not encompass the test set on the lower end for Chapter 3 and 4 while on the higher end for Chapter 5 according to maximum/minimum values. Therefore, it could be expected that this discrepancy in the value range of training and test sets would lead to uncertainty in modeling as the variability of the test set is not completely involved in the training set. In response to the variability in $\Psi_{stem}$. It was observed that statistically important wavelengths were distributed in VIS, NIR, and SWIR over the entire spectrum in hyperspectral measurements (Chapter 3). The spectral information related to $\Psi_{stem}$ variation, is explored when raw reflectance was processed by various transformations (raw reflectance, 1D, 2D, CR, and SI) and variable selection (Spearman correlation and RFECV) approaches. The most accurate model uses SI as input (Table 3.7) indicating that SI transformation is able to largely increase the variance of inputs, based on the entire spectrum through the combination of pair-wise bands. The RMSE (110 kPa) of this model is smaller than half of the standard deviation of the test set (139.5 kPa is half of the standard deviation of the test set), which indicates it is a satisfactory prediction (Singh et al., 2005). Its RPIQ is equal to 4.28, which is considered excellent according to the study of Efron (2014). The second accurate and the third accurate model uses raw reflectance and the first derivative as input, respectively. PLSR subsequently integrates the shared variance to reach the maximum potential of $\Psi_{stem}$ correlation. When breaking down the important spectral regions identified by variable selection approaches for each transformation group, it shows that the significant wavelengths are assembled around blue, red, and red edge bands in VIS, two weak water absorption bands at 970 and 1200 nm in NIR, and two strong absorption bands at 1400 and 1940 nm with dry matter-related bands in SWIR. When applying these relationships to estimate GWS at a larger scale, measurements must be taken by remote, rather than proximal sensors. These sensors would be mounted on drones and satellites, due to the advances in their availability and accessibility. When applying aerial-based or space-based observations, water absorption bands, become redundant as solar radiation is mostly absorbed by water vapor in the atmosphere, before attaining the objects. Therefore, it does not appear in most multispectral sensors with a sufficient spatial resolution (<0.5 m) (Maimaitijiang et al., 2020), and blue, red, and red edge bands are the main contributors to detecting moderate water stress in airborne or spaceborne multispectral platforms. This was shown in Chapter 4, which demonstrates that TCARI, ExG, and TCARI/OSAVI, computed from blue, red, and red edge bands, outperformed GNDVI, NDRE, and NDVI, which includes NIR in VI computation. This implies leaf pigments, especially chlorophylls, can better depict the variation in moderate water deficit for Pinot Noir than the alteration of canopy geometry and cell structure does. The better performance of TCARI compared to the

modified index (TCARI/OSAVI) in capturing $\Psi_{stem}$ variability also shows LAI of Pinot Noir during mild dehydration is relatively unchanged. Given these results, it is recommended to use VIs including blue, red, and red edge bands to estimate $\Psi_{stem}$ under mild dehydration regimes for Pinot Noir.

### *6.1.2 Interoperation among Data Collected from Multiple Sources*

As SWIR and thermal infrared band, the main source of robust indicators for assessing vegetation water status, is missing in most of the commercial satellite and drone multispectral sensors (Jang et al., 2020; Van Beek, 2018), it is anticipated that the practical application of spaceborne or airborne multispectral for irrigation scheduling is limited. One solution to address this is to take advantage of information from other data sources.

In Chapter 3, ASD hyperspectral spectroradiometer explores the relationships between spectral information and changes in $\Psi_{stem}$. However, an ASD spectroradiometer provides only point measurements without spatial information, and airborne images can sense fields in high spatial resolution (under 5 cm) at once. By calibrating the UAV-based pixels with $\Psi_{stem}$, the spatial GWS across vineyards can be estimated by extrapolating the relative difference of spectral value between pixels of the drone images. Available hyperspectral airborne sensors have not been cost-effective at the moment to meet the needs of irrigation management. It would be natural to find a way to improve the sensing capacities of commercially available multispectral sensors for irrigation scheduling, despite their spectral resolution often being limited to three to five bands in VIS-NIR regions.

The result of the best correlated multispectral VI with $\Psi_{stem}$ ($R^2 = 0.35$) in Chapter 4 is less than satisfactory, which can be linked to the goal of this study – detecting moderate water stress. The dehydration level is not severe or long enough to affect canopy structure and leaf pigments, so not causing significant variations in multispectral reflectance. To improve the sensing capabilities in estimating GWS, additional data sources, providing environmental information, are employed to help depict the dynamics. The roles that soil, plant, and atmosphere play in crop water consumption are taken into account. To make the relationships between $\Psi_{stem}$ variability and environmental variables simple, individual differences originating from cultivar, age, and cultivation practices were minimized by choosing the monocultivar, Pinot Noir, with a similar age (planted in 1998-2000) experiencing the same vineyard management in the trials. Besides, given the mild temperature and rainfall in the study area over the study period, and Pinot Noir being near-anisohydric (Gutiérrez-Gamboa et al., 2019), the stomatal closure forced by soil water deficit may be eliminated. It is assumed the GWS variation in this study is an integrative response to atmospheric supply and demand, canopy ET, and soil water availability during the growing season. The results show that $\Psi_{stem}$ detection can reach the accuracy of RMSE of 138 kPa with RPIQ of 2.75 when employing DOY, elevation, slope, $EC_a$, and daily average windspeed together with UAV-based EVI. However, either TCARI-based or ExG-based model does not provide satisfactory prediction because the RMSE value of TCARI-based or ExG-based model (146 and 138 kPa, respectively) appears to be larger than half of the standard deviation of the test set (134 kPa is half of the standard deviation of the test set) (Singh et al., 2005). The RPIQ value of TCARI-based or ExG-based model (2.6 and 2.75, respectively) is considered excellent according to the study of Efron (2014).

Modeling based on environmental variables only, without including multispectral VI was tested as well to assess if it is necessary to use VI as core inputs. It turns out that the best modeling result (RMSE = 145 kPa) without using VI does not outperform the one with VI (RMSE = 138 kPa).

To facilitate irrigation management in vineyards, SHAP was used to interpret the relationships, contribution weight, and directionality, between the selected environmental variables and $\Psi_{stem}$. Hierarchical clustering was implemented to reduce collinearity among input variables while assisting in generating reasonable causal relationships for an easier interpretation. The reasonable causal relationships were determined according to whether the contribution of a predictor to the variation of $\Psi_{stem}$ could be supported by the available literature. Two interesting phenomena were revealed during the relationship interpretation. Vines at lower elevations do not usually result in a canopy with higher water status due to water accumulation. It is deduced these higher and older terraces in NZ were covered with fine loess, so the developing soil has a property of higher water-holding capacity, leading to more hydrated grapevines. Another phenomenon is that higher wind speed does not usually lead to a higher level of water stress in grapevines induced by ET, which may be associated with reducing water loss via stomatal closure or less reliance on evaporative cooling.

### *6.1.3 Upscaling of $\Psi_{stem}$ Monitoring*

From a practical perspective, satellite-based sensing is more cost-effective than UAV-based sensing for irrigation scheduling in terms of coverage area per time requirement for sensing, effort in post-processing, radiometric quality, and regular acquisition (Borgogno-Mondino et al., 2018; Matese & Di Gennaro, 2015). However, there are two constraints when applying satellite images to estimating GWS. Satellite sensors with low to moderate resolution, which are usually freely accessible and more of practical use for growers, can not provide reliable estimation when sensing over vertical-oriented vineyards, due to interference resulting from inter-row vegetation or bare soil (Khaliq et al., 2019). The other limitation is the time required for post-processing and thus the satellite image products are not delivered in real time (Gautam & Pagay, 2020; Van Beek, 2018). In addition, weather parameters, clouds, and fog, often have an impact over the usability of satellite applications for crop monitoring.

Two-stage calibration and a prediction tool are presented to address the issue of restricted spatial resolution and inaccessible images (Chapter 5), by taking advantage of the $\Psi_{stem}$-related spectral features (Chapter 3) and important environmental variables (Chapter 4). Two-stage calibration first calibrates UAV images with ground measurements, then calibrates satellite images with calibrated UAV data along with removing the information contributed by the inter-row elements. Note that either the model for calibrating UAV images acquired in 2020/2021 or 2021/2022 provides a satisfactory prediction because the RMSE of the 2020/2021 model and 2021/2022 model is equal to 113 and 106 kPa, respectively. It is smaller than half of the standard deviation of the test set of 2020/2021 (138.5 kPa is half of the standard deviation of the 2020/2021 test set) and 2021/2022 (115 kPa is half of the standard deviation of the 2021/2022 test set) (Singh et al., 2005). In addition, the RPIQ of the 2020/2021 model and 2021/2022 model is 4.2 and 2.2, respectively. They are classified as excellent and very good according to the study of Efron (2014). This enables $\Psi_{stem}$ to

be approximated (RMSE of 59 kPa at the second calibration in the first growing season) at a larger scale with more accurate and site-specific satellite images. The prediction model is developed based on the data obtained in the first growing season with RMSE of 31 kPa, and it performs well when applied to predict the $\Psi_{stem}$ variation in the second growing season (r of 0.89 at Pencarrow and r of 0.87 at Wharekauhou). This model is established based on DOY, total daily rainfall, total daily PET, irrigation, fertigation, plucking, trimming accounting for temporal variation, NDVI, $EC_a$, elevation, and slope accounting for spatial variation. The major advantage of this model is that it allows one to get an idea of how GWS will proceed during the next 30 days on a large scale based on the quality of the weather forecast. However, there are two main weaknesses in the prediction model, so the model can not be applied to viticultural management directly. At this stage, the outputs of the prediction model can not be viewed as predicted $\Psi_{stem}$ values as they only present the temporal trend of $\Psi_{stem}$ changes for the vineyard. The model needs calibration at the beginning of the growing season to truly represent the changes in $\Psi_{stem}$. In addition, the model only outputs the mean $\Psi_{stem}$ for the entire vineyard instead of generating the spatial map of $\Psi_{stem}$ over the field. It is assumed that by using weather parameters recorded by a wireless sensor network spatially and temporally, spatio-temporal variation can be added to the prediction model to further provide a spatial map of $\Psi_{stem}$ across vineyards along the growing season.

### 6.2 Conclusion

In the context of more severe projected climate change, more restricted environmental regulations, and growing competition in international markets, it becomes of utmost importance to produce grapes, within every block, reaching high-quality standards for intended products. Meeting the market's demand is more profitable, especially for high-class wines, than selling downgraded wines whose source has a fraction of low-quality grapes, as low-quality attributes, even a small proportion, prevail in the wines. This has led to a radical review of viticultural management and emerging techniques, with the goal of satisfying the real needs of grapevines to maximize and stabilize quality, while attaining a more efficient use of inputs to reduce operation costs against the increasing trend.

The provision of georeferenced spatial datasets provides growers, with opportunities to optimize the processes of decision-making in controlling GWS to reach improved grape/wine quality by using several site-specific cultivation practices (irrigation, trimming, plucking, mulching) to assist in regulated DI. Meanwhile, savings are obtained on input costs (decreasing water wastage, harvesting, and pumping) and labor costs because excess growth is suppressed with a reduced need for pruning. This study provides a step toward the practical application of multispectral satellite images for precision irrigation management. The results reveal that (i) spectral features related to $\Psi_{stem}$ variation of Pinot Noir are distributed over VIS (400–430 and 650–750 nm), NIS (800–1250 nm), and SWIR (1370–1420, 1500–1595, 1700–1720, 1850–1890, and 2050–2370 nm) of the spectrum. (ii) Environmental variables (temporal trend, soil/terrain characteristics, and weather parameters) are important estimators that can assist multispectral images in describing $\Psi_{stem}$ variation. Either TCARI together with DOY and elevation or ExG together with DOY, elevation, weekly mean wind speed, and $EC_a$, is useful in describing $\Psi_{stem}$ changes of Pinot Noir between flowering and

veraison in Martinborough, NZ. (iii) It is feasible to develop a $\Psi_{stem}$ prediction model that can provide mean $\Psi_{stem}$ estimation 30 days in advance for vineyards. These findings increase the knowledge of estimating GWS remotely with practical solutions for in-field application, and they pave the way for developing automatic data-driven decision support systems that produce grapes with higher quality. From a broad perspective, the improvement of the application of satellite images not only can be applied to irrigation scheduling but also to other viticultural monitoring targets including disease/pest monitoring, nutrient status assessment, and vigor/leaf area evaluation.

### 6.3 Limitation

Due to the research goal (monitoring early to moderate water stress) of this study, the sampling window of $\Psi_{stem}$ was concentrated during the two months before grapevines enter the veraison period as it is the most critical timeframe for controlling GWS and determining grape quality. It is expected that the variability of environmental variables, including weather parameters and soil moisture availability, is small, so the model does not hold the same level of accuracy when performing in other areas with different climate patterns. In addition, due to limited manpower and the number of pressure bombs, the sampling capability of $\Psi_{stem}$ is constrained, together with a restricted range of inputs from the aspects of site, soil/terrain characteristics, grapevine cultivars, canopy age, cultivation practices, and growing seasons. Thus, the developed models are not ready to be transferred into other situations, but they established the concept for applying RS information, acquired from both UAV and satellite platforms to irrigation scheduling, at a practical level. In addition, the developed models are not able to generate spatial GWS maps at this stage. There is one more step needed for map creation for each developed model developed in Chapters 4 and 5, respectively. In Chapter 4, it is required to collect the coordinates of the trunk of each grapevine within vineyards because the individual canopy is not distinguishable in the UAV images. The Chapter 4-derived model is computed based on the spectral information of the pure pixels of an individual canopy, so the spatial GWS maps can be produced after each canopy, within a given block, is identified. The Chapter 5-derived model currently shows only temporal variability of the mean $\Psi_{stem}$ of the vineyard. A variable that can approximate the spatio-temporal variation of $\Psi_{stem}$ needs to be involved in the modeling process, and the prediction model is then capable of providing spatial GWS prediction maps.

### 6.4 Future research

The satellite-based models developed by this study will enable temporal monitoring of GWS within the critical period, to assist in identifying management blocks corresponding to areas with the same level of water stress. Nevertheless, prior to the practical use of developed models in decision support for irrigation management in vineyards, some further research is required. Future work should be directed to (i) the development of an integrative online service: Focus on creating a comprehensive online platform that seamlessly integrates remote sensing data and tools. This platform should provide easy access for users to analyze and interpret data for various applications.

(ii) the assessment of cultivation practices for adjusting GWS: Investigate how cultivation practices can be optimized to better suit the requirements of precision agriculture, specifically focusing on adjusting the Grape Water Stress (GWS) levels to enhance grapevine health and production.
(iii) the amount of irrigated water needed: Determine the precise amount of irrigation water needed in vineyards, taking into account factors such as soil moisture, weather conditions, and plant stress levels. This information can contribute to more efficient water management strategies.
(iv) the evaluation of RS information collected in other forms or from other platforms: Explore the utility of remote sensing data collected through different means or from alternative platforms beyond the current sources. Assess how these alternative data sources can complement or enhance existing information.
(v) the assessment of overall benefits and costs: Conduct a thorough assessment of the overall benefits and costs associated with implementing remote sensing techniques in viticulture. This analysis should consider both short-term gains, such as increased quality and reduced resource usage, and long-term benefits, including improved sustainability and economic viability.

To provide an easy-to-approach service, it is necessary to develop a one-step application, either web-based or app-based, to allow end-users or viticulturists to access spatial GWS information by specifying the coordinates of vineyards. To this end, future work should consider the integration of satellite image and environmental information gathering, automatic pre-processing, statistical analysis, cloud computing, and front-end rendering for building up online decision support systems. This will broaden the application and improve the usability of satellite images in viticulture.
After acquiring the spatial GWS maps, agricultural practices should be assessed at the vineyard level in terms of their efficiency and cost in standardizing GWS across fields and attaining intended quality. These practices include canopy management (thinning, adjusting trellis system height, pruning, adopting new training systems) which determines the water consumption by controlling the amount of intercepted sunlight and the microclimate inside the canopy. The application of mulches can influence soil moisture retention by reducing soil evaporation. The use of exogenous compounds and shadow nets would reduce canopy ET. Moreover, variable rate irrigation, implemented by mobile sprayer or zoned drip irrigation, should be evaluated as a supplementary measure for its feasibility in GWS management.

In addition, the quantity of water that should be applied when there is an indication of irrigation, should be better understood. This requires the plant-based models to work with other measurements, such as soil moisture balance methods, to quantify the optimal water demands.

Some satellite platforms offer SWIR information or microwave information at high spatial resolution, including WorldView-3 (with 8 bands over 1195 to 2365 nm at 3.7 m resolution) and PAZ (with X-band over 2.8 to 5.2 cm at 0.22 x 0.59 m resolution). The sufficient resolution of these systems could provide insights into crop water content and support decision-making in GWS control. Future work can focus on comparing the applicability of different RS data sources with that of VIS-NIR information, while evaluating their usability in establishing GWS prediction models.
Lastly, some work should be undertaken on pluriannual monitoring to assess the marginal revenue when including the precision irrigation stack into different types of

viticultural management systems at various terroirs or regions. It is necessary to consider holistically from the following perspectives: the RS information types and acquisition platforms, spatial data processing chains, visualization products, corresponding cultivation practices, labor costs, variable rate irrigation devices, long-term environmental policies, and water scarcity. Subsequently, the applications of precision irrigation can be justified and make contributions toward the long-term benefits of the viticultural industry.

**Reference**

Acevedo-Opazo, C., Tisseyre, B., Guillaume, S., & Ojeda, H. (2008). The potential of high spatial resolution information to define within-vineyard zones related to vine water status. *Precision Agriculture, 9*(5), 285-302. https://doi.org/10.1007/s11119-008-9073-1

Baeza, P., Junquera, P., Peiro, E., Lissarrague, J. R., Uriarte, D., & Vilanova, M. (2019). Effects of vine water status on yield components, vegetative response and must and wine composition. *Advances in Grape and Wine Biotechnology*.

Borgogno-Mondino, E., Lessio, A., Tarricone, L., Novello, V., & De Palma, L. (2018). A comparison between multispectral aerial and satellite imagery in precision viticulture. *Precision Agriculture, 19*(2), 195-217.

Choné, X. (2001). Stem Water Potential is a Sensitive Indicator of Grapevine Water Status. *Annals of Botany, 87*(4), 477-483. https://doi.org/10.1006/anbo.2000.1361

Efron, B. (2014). Estimation and accuracy after model selection. *Journal of the American Statistical Association*, *109*(507), 991-1007.

Fernandes-Silva, A., Oliveira, M., Paço, T. A., & Ferreira, I. (2019). Deficit irrigation in Mediterranean fruit trees and grapevines: Water stress indicators and crop responses. *Irrigation in Agroecosystems. London, United Kingdom*.

Gautam, D., & Pagay, V. (2020). A review of current and potential applications of remote sensing to study the water status of horticultural crops. *Agronomy, 10*(1), 140.

Gutiérrez-Gamboa, G., Pérez-Donoso, A. G., Pou-Mir, A., Acevedo-Opazo, C., & Valdés-Góme, H. (2019). Hydric behaviour and gas exchange in different grapevine varieties (Vitis vinifera L.) from the Maule Valley (Chile). *South African Journal of Enology and Viticulture, 40*(2), 1-1.

IPCC. (2014). *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change*.

Jang, G., Kim, J., Yu, J.-K., Kim, H.-J., Kim, Y., Kim, D.-W., Kim, K.-H., Lee, C. W., & Chung, Y. S. (2020). Cost-effective unmanned aerial vehicle (UAV) platform for field plant breeding application. *Remote Sensing, 12*(6), 998.

Khaliq, A., Comba, L., Biglia, A., Ricauda Aimonino, D., Chiaberge, M., & Gay, P. (2019). Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. *Remote Sensing, 11*(4), 436.

Maimaitijiang, M., Sagan, V., Sidike, P., Daloye, A. M., Erkbol, H., & Fritschi, F. B. (2020).

Crop monitoring using satellite/UAV data fusion and machine learning. *Remote Sensing, 12*(9), 1357.

Matese, A., & Di Gennaro, S. F. (2015). Technology in precision viticulture: A state of the art review. *International journal of wine research, 7*, 69-81.

Singh, J., Knapp, H. V., Arnold, J. G., & Demissie, M. (2005). Hydrological modeling of the Iroquois river watershed using HSPF and SWAT 1. *JAWRA Journal of the American Water Resources Association*, *41*(2), 343-360.

Solomon, S., Qin, D., Manning, M., Averyt, K., & Marquis, M. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge university press.

Stocker, T. (2014). *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge university press.

Van Beek, J. (2018). *High Spatial Resolution Satellite Imagery for Irrigation Scheduling in Hedgerow Cropping Systems* KU Leuven].

## Appendices

Appendix 1: In-situ measurements of grapevine water status for the two study vineyards over two study years. Psi is pounds per square inch.

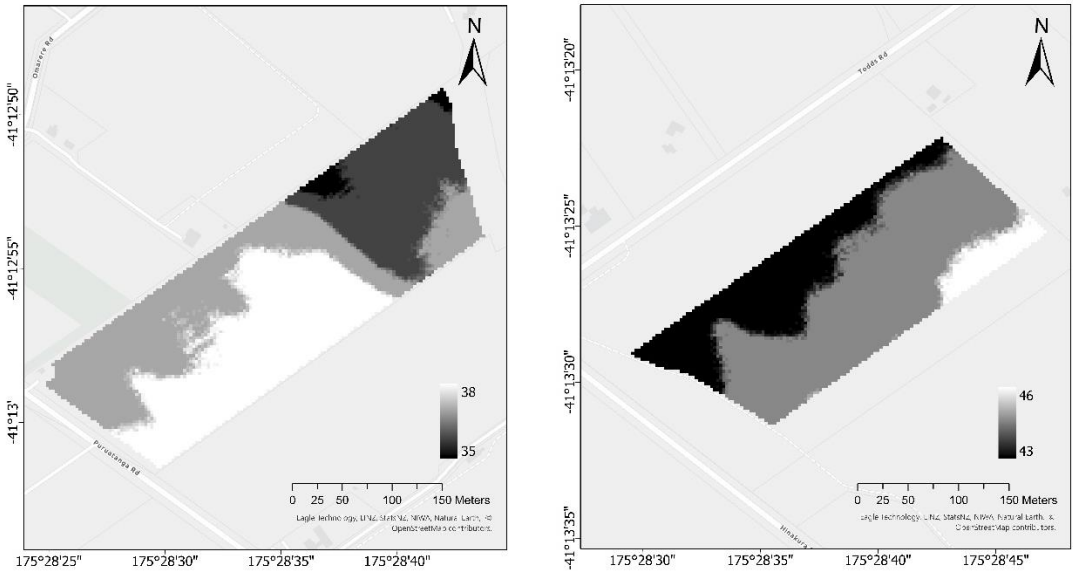
| Measurement date | Vineyard | First leaf (Psi) | Second leaf (Psi) |
|---|---|---|---|
| 20201204 | Pencarrow | 82 | 65 |
| 20201204 | Pencarrow | 90 | 50 |
| 20201204 | Pencarrow | 50 | 60 |
| 20201204 | Pencarrow | 65 | 60 |
| 20201204 | Pencarrow | 70 | 73 |
| 20201204 | Pencarrow | 65 | 95 |
| 20201204 | Pencarrow | 45 | 70 |
| 20201204 | Pencarrow | 50 | 55 |
| 20201204 | Pencarrow | 55 | 87 |
| 20201204 | Pencarrow | 40 | 60 |
| 20210114 | Pencarrow | 155 | 145 |
| 20210114 | Pencarrow | 115 | 135 |
| 20210114 | Pencarrow | 150 | 100 |
| 20210114 | Pencarrow | 107.5 | 95 |
| 20210114 | Pencarrow | 105 | 115 |
| 20210114 | Pencarrow | 137.5 | 175 |
| 20210114 | Pencarrow | 92.5 | 90 |
| 20210114 | Pencarrow | 127.5 | 110 |
| 20210114 | Pencarrow | 115 | 100 |
| 20210114 | Pencarrow | 90 | 125 |
| 20210114 | Pencarrow | 95 | 155 |
| 20210122 | Pencarrow | 127.5 | 125 |
| 20210122 | Pencarrow | 135 | 82.5 |
| 20210122 | Pencarrow | 100 | 100 |
| 20210122 | Pencarrow | 115 | 110 |
| 20210122 | Pencarrow | 120 | 130 |
| 20210122 | Pencarrow | 115 | 125 |
| 20210122 | Pencarrow | 143.5 | 115 |
| 20210122 | Pencarrow | 145 | 150 |
| 20210122 | Pencarrow | 97.5 | 102.5 |
| 20210122 | Pencarrow | 145 | 142.5 |
| 20210122 | Pencarrow | 105 | 100 |
| 20210201 | Pencarrow | 145 | 127.5 |
| 20210201 | Pencarrow | 112.5 | 125 |
| 20210201 | Pencarrow | 137.5 | 120 |
| 20210201 | Pencarrow | 130 | 137.5 |
| 20210201 | Pencarrow | 147.5 | 170 |

| Measurement date | Vineyard | First leaf (Psi) | Second leaf (Psi) |
|---|---|---|---|
| 20210201 | Pencarrow | 145 | 162.5 |
| 20210201 | Pencarrow | 185 | 200 |
| 20210201 | Pencarrow | 160 | 145 |
| 20210201 | Pencarrow | 120 | 107.5 |
| 20210201 | Pencarrow | 180 | 160 |
| 20210201 | Pencarrow | 110 | 105 |
| 20210201 | Pencarrow | 150 | 145 |
| 20210201 | Pencarrow | 75 | 65 |
| 20210201 | Pencarrow | 155 | 152.5 |
| 20210201 | Pencarrow | 125 | 125 |
| 20210201 | Pencarrow | 165 | 130 |
| 20210201 | Pencarrow | 185 | 180 |
| 20210201 | Pencarrow | 162.5 | 140 |
| 20201127 | Wharekauhau | 80 | 70 |
| 20201127 | Wharekauhau | 75 | 60 |
| 20201127 | Wharekauhau | 55 | 52.5 |
| 20201127 | Wharekauhau | 67.5 | 43.5 |
| 20201127 | Wharekauhau | 65 | 65 |
| 20201127 | Wharekauhau | 50 | - |
| 20201127 | Wharekauhau | 55 | 40 |
| 20201127 | Wharekauhau | 60 | 55 |
| 20201127 | Wharekauhau | 60 | 50 |
| 20201127 | Wharekauhau | 50 | 55 |
| 20201127 | Wharekauhau | 45 | - |
| 20201204 | Wharekauhau | 70 | 80 |
| 20201204 | Wharekauhau | 82.5 | 55 |
| 20201204 | Wharekauhau | 55 | 55 |
| 20201204 | Wharekauhau | 75 | 70 |
| 20201204 | Wharekauhau | 85 | 62 |
| 20201204 | Wharekauhau | 78 | 45 |
| 20201204 | Wharekauhau | 80 | 60 |
| 20201204 | Wharekauhau | 70 | 43 |
| 20210114 | Wharekauhau | 175 | 165 |
| 20210114 | Wharekauhau | 125 | 115 |
| 20210114 | Wharekauhau | 145 | 145 |
| 20210114 | Wharekauhau | 140 | 147.5 |
| 20210114 | Wharekauhau | 165 | 160 |
| 20210114 | Wharekauhau | 110 | 107.5 |
| 20210114 | Wharekauhau | 145 | 150 |
| 20210114 | Wharekauhau | 97.5 | 140 |
| 20210122 | Wharekauhau | 140 | 115 |

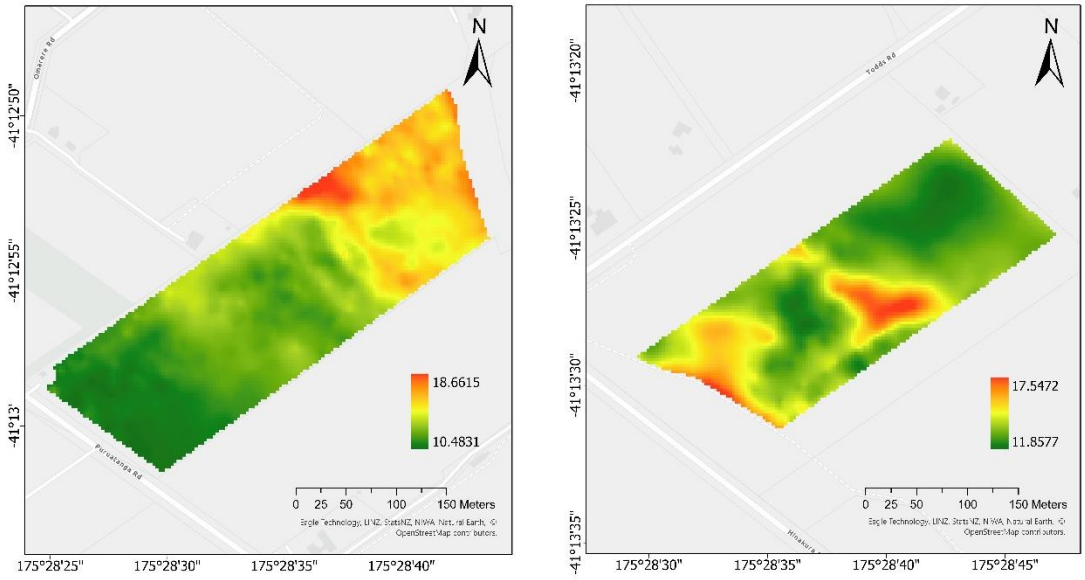
| Measurement date | Vineyard | First leaf (Psi) | Second leaf (Psi) |
|---|---|---|---|
| 20210122 | Wharekauhau | 137.5 | 135 |
| 20210122 | Wharekauhau | 130 | 120 |
| 20210122 | Wharekauhau | 210 | 180 |
| 20210122 | Wharekauhau | 150 | 165 |
| 20210122 | Wharekauhau | 130 | 115 |
| 20210122 | Wharekauhau | 157.5 | 155 |
| 20210122 | Wharekauhau | 160 | 155 |
| 20211129 | Pencarrow | 55 | 90 |
| 20211129 | Pencarrow | 110 | 80 |
| 20211129 | Pencarrow | 85 | 57.5 |
| 20211129 | Pencarrow | 75 | 95 |
| 20211129 | Pencarrow | 95 | 70 |
| 20211129 | Pencarrow | 85 | 60 |
| 20211129 | Pencarrow | 150 | 90 |
| 20211129 | Pencarrow | 55 | 75 |
| 20211129 | Pencarrow | 70 | 130 |
| 20211209 | Pencarrow | 45 | 55 |
| 20211209 | Pencarrow | 50 | 55 |
| 20211209 | Pencarrow | 60 | 80 |
| 20211209 | Pencarrow | 60 | 50 |
| 20211209 | Pencarrow | 55 | 50 |
| 20211209 | Pencarrow | 87.5 | 37.5 |
| 20211209 | Pencarrow | 67.5 | 40 |
| 20211209 | Pencarrow | 75 | 80 |
| 20211209 | Pencarrow | 50 | 47.5 |
| 20220111 | Pencarrow | 70 | 90 |
| 20220111 | Pencarrow | 162.5 | 92.5 |
| 20220111 | Pencarrow | 95 | 50 |
| 20220111 | Pencarrow | 90 | 95 |
| 20220111 | Pencarrow | 130 | 125 |
| 20220111 | Pencarrow | 160 | 145 |
| 20220111 | Pencarrow | 70 | 100 |
| 20220111 | Pencarrow | 60 | 50 |
| 20220111 | Pencarrow | 80 | 90 |
| 20220121 | Pencarrow | 165 | 147.5 |
| 20220121 | Pencarrow | 120 | 100 |
| 20220121 | Pencarrow | 95 | 100 |
| 20220121 | Pencarrow | 50 | 47.5 |
| 20220121 | Pencarrow | 127.5 | 127.5 |
| 20220121 | Pencarrow | 140 | 145 |
| 20220121 | Pencarrow | 80 | 85 |

| Measurement date | Vineyard | First leaf (Psi) | Second leaf (Psi) |
|---|---|---|---|
| 20220121 | Pencarrow | 145 | 107.5 |
| 20220121 | Pencarrow | 150 | 165 |
| 20211129 | Wharekauhau | 80 | 90 |
| 20211129 | Wharekauhau | 85 | 75 |
| 20211129 | Wharekauhau | 40 | 100 |
| 20211129 | Wharekauhau | 100 | 60 |
| 20211129 | Wharekauhau | 80 | 85 |
| 20211129 | Wharekauhau | 90 | 70 |
| 20211129 | Wharekauhau | 25 | 65 |
| 20211129 | Wharekauhau | 100 | 85 |
| 20211129 | Wharekauhau | 82.5 | 65 |
| 20211209 | Wharekauhau | 45 | 70 |
| 20211209 | Wharekauhau | 75 | 50 |
| 20211209 | Wharekauhau | 55 | 105 |
| 20211209 | Wharekauhau | 25 | 60 |
| 20211209 | Wharekauhau | 60 | 80 |
| 20211209 | Wharekauhau | 50 | 115 |
| 20211209 | Wharekauhau | 65 | 50 |
| 20211209 | Wharekauhau | 70 | 57.5 |
| 20211209 | Wharekauhau | 55 | 95 |
| 20220111 | Wharekauhau | - | 90 |
| 20220111 | Wharekauhau | 80 | 95 |
| 20220111 | Wharekauhau | 140 | 110 |
| 20220111 | Wharekauhau | 125 | 100 |
| 20220111 | Wharekauhau | 85 | 150 |
| 20220111 | Wharekauhau | 115 | 115 |
| 20220111 | Wharekauhau | 135 | 115 |
| 20220111 | Wharekauhau | 105 | 97.5 |
| 20220111 | Wharekauhau | 110 | 125 |
| 20220121 | Wharekauhau | 145 | 145 |
| 20220121 | Wharekauhau | 142.5 | 120 |
| 20220121 | Wharekauhau | 105 | 157.5 |
| 20220121 | Wharekauhau | 102.5 | 142.5 |
| 20220121 | Wharekauhau | 115 | 115 |
| 20220121 | Wharekauhau | 160 | 147.5 |
| 20220121 | Wharekauhau | 147.5 | 165 |
| 20220121 | Wharekauhau | 137.5 | 135 |
| 20220121 | Wharekauhau | 150 | 165 |

Appendix 2: Elevation maps, derived from the digital elevation model, for the two study vineyards (the left is Wharekauhau and the right is Pencarrow).



Appendix 3: Apparent electrical conductivity maps generated by interpolating EM38-based data for the two study vineyards (the left is Wharekauhau and the right is Pencarrow).



Appendix 4: Python scripts that are developed in Chapter 3-5.

GitHub: https://github.com/HsiangEnWei/PhD-thesis/branches/all

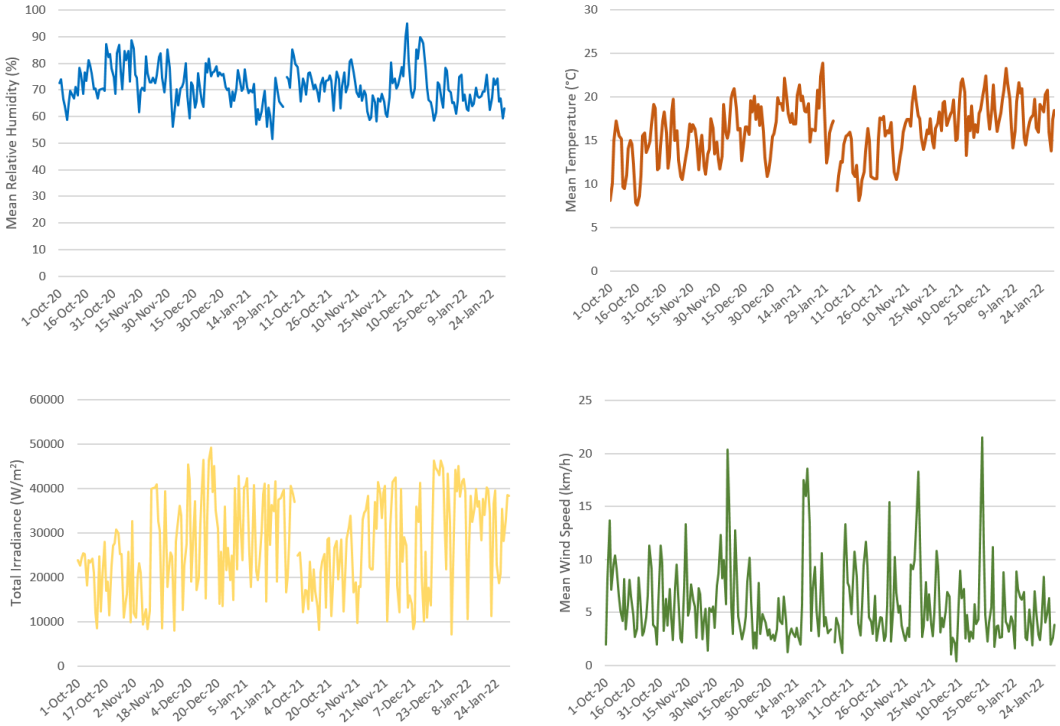
Appendix 5: Descriptive statistics for $\Psi_{stem}$ dataset in Chapter 3-5.

| | | Sample size | Mean (kPa) | Standard deviation (kPa) | Maximum value (kPa) | Minimum value (kPa) |
|---|---|---|---|---|---|---|
| Chapter 3 | Training set | 59 | 751 | 273 | 1344 | 328 |
| | Test set | 26 | 754 | 279 | 1258 | 310 |
| Chapter 4 | Training set | 103 | 699 | 263 | 1344 | 310 |
| | Test set | 45 | 701 | 268 | 1327 | 293 |
| Chapter 5 | 2020/2021 Training set | 59 | 751 | 274 | 1327 | 310 |
| | 2020/2021 Test set | 26 | 754 | 277 | 1344 | 345 |
| | 2021/2022 Training set | 44 | 629 | 231 | 1077 | 293 |
| | 2021/2022 Test set | 19 | 627 | 230 | 1086 | 345 |

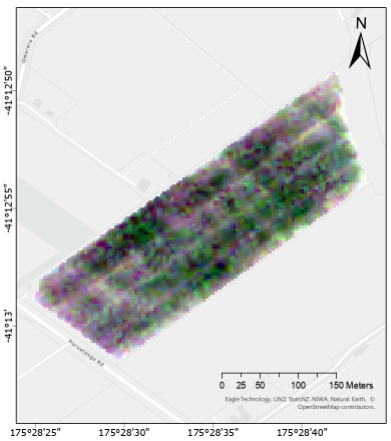
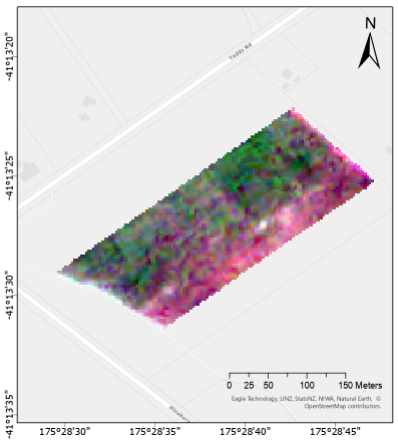
Appendix 6: Results of modeling performance using vegetation indices as inputs in Chapter 3.

| Vegetation Indices | $R^2$ | RMSE |
|---|---|---|
| NDWI | 0.03 | 271 |
| SRWI | 0.03 | 271 |
| MSI | 0.06 | 267 |
| MDWI | 0.03 | 271 |
| SR | 0.18 | 248 |
| DDI | 0.16 | 252 |
| FWBI | 0.08 | 263 |
| SWIWSI | 0.08 | 264 |
| WI | 0.09 | 263 |
| WABI | 0.4 | 212 |
| PRI | 0.41 | 210 |
| NDVI | 0.26 | 236 |

Appendix 7: Weather data recorded by the on-site weather station (175.4741, -41.2247 WGS84) and established by HARVEST.com (http://harvest.com/) for the two study periods.



Appendix 8: Examples of UAV images and satellite images used in this study. All images are taken on 14 January 2021 and displayed in RGB.

| | Vineyard | |
|---|---|---|
| Sensing platform | Wharekauhau | Pencarrow |
| UAV (Multispectral Phantom 4 ) |   |   |
| Satellite (PlanetScope) |   |   |

Appendix 9: Analysis of multicollinearity in the dataset used in Chapter 5.

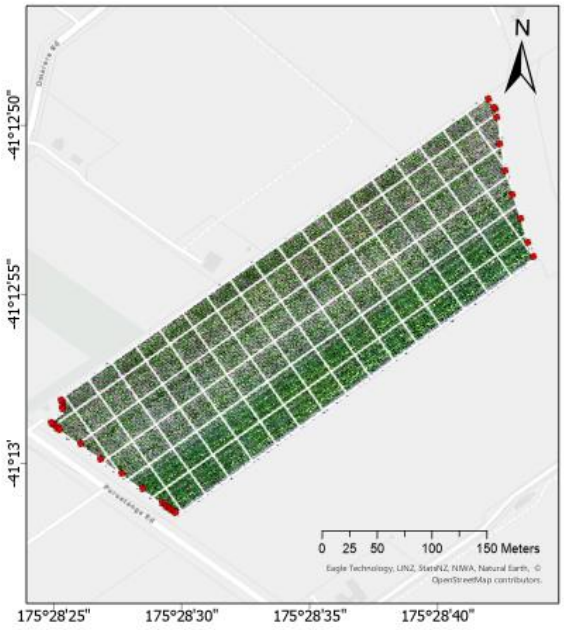
1. Correlation matrix between vegetation indices

| | NDVI | GNDVI | SR | MSAVI | OSAVI | R/G index | ExG | EVI | TCARI | MCARI | MTVI1 | NDRE | NGRDI | CL rededge | DVI | VARI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NDVI | 1 | | | | | | | | | | | | | | | |
| GNDVI | 0.80 | 1 | | | | | | | | | | | | | | |
| SR | 0.96 | 0.88 | 1 | | | | | | | | | | | | | |
| MSAVI | 0.59 | 0.20 | 0.47 | 1 | | | | | | | | | | | | |
| OSAVI | 0.73 | 0.34 | 0.62 | 0.98 | 1 | | | | | | | | | | | |
| R/G index | -0.88 | -0.42 | -0.77 | -0.73 | -0.83 | 1 | | | | | | | | | | |
| ExG | 0.41 | -0.16 | 0.24 | 0.86 | 0.84 | -0.72 | 1 | | | | | | | | | |
| EVI | 0.55 | 0.15 | 0.43 | 1.00 | 0.97 | -0.71 | 0.86 | 1 | | | | | | | | |
| TCARI | -0.53 | -0.83 | -0.70 | 0.25 | 0.10 | 0.17 | 0.47 | 0.30 | 1 | | | | | | | |
| MCARI | 0.80 | 0.42 | 0.72 | 0.92 | 0.96 | -0.87 | 0.78 | 0.90 | -0.09 | 1 | | | | | | |
| MTVI1 | 0.54 | 0.11 | 0.41 | 1.00 | 0.97 | -0.72 | 0.90 | 1.00 | 0.32 | 0.897729 | 1 | | | | | |
| NDRE | 0.48 | 0.75 | 0.59 | 0.00 | 0.09 | -0.14 | -0.30 | -0.04 | -0.63 | 0.06 | -0.07 | 1 | | | | |
| NGRDI | 0.89 | 0.45 | 0.80 | 0.72 | 0.82 | -1.00 | 0.71 | 0.70 | -0.21 | 0.88 | 0.71 | 0.16 | 1 | | | |
| CL rededg | 0.45 | 0.71 | 0.56 | -0.03 | 0.06 | -0.13 | -0.29 | -0.07 | -0.61 | 0.03 | -0.09 | 0.99 | 0.15 | 1 | | |
| DVI | 0.55 | 0.16 | 0.43 | 1.00 | 0.97 | -0.70 | 0.87 | 1.00 | 0.29 | 0.90 | 1.00 | -0.03 | 0.69 | -0.06 | 1 | |
| VARI | 0.87 | 0.45 | 0.79 | 0.73 | 0.82 | -0.96 | 0.65 | 0.72 | -0.19 | 0.87 | 0.72 | 0.17 | 0.96 | 0.15 | 0.71 | |

2. Pearson correlation coefficient between rainfall and potential evapotranspiration data = -0.34

Appendix 10: Sampling grids used for systematic sampling over the two study vineyards. The red dots are the ground control points.

Wharekauhau




Pencarrow