Assessing Field Spectroscopic Methods for Grapevine Chlorophyll Content Estimation

by

Diana Parton
B.Sc., University of Victoria, 2007

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Geography

Supervisory Committee

Assessing Field Spectroscopic Methods for Grapevine Chlorophyll Content Estimation

by

Diana Parton
B.Sc., University of Victoria, 2007

Supervisory Committee

Dr. K. Olaf Niemann (Department of Geography)
Supervisor

Dr. Mark S. Flaherty, (Department of Geography)
Departmental Member

Abstract

Supervisory Committee
Dr. K. Olaf Niemann
Supervisor
Dr. Mark Flaherty
Departmental Member

Vancouver Island, British Columbia, is at the northern extent of natural climate zones conducive for grape growing, making vineyards susceptible to any changing weather patterns and temperature extremes. Grapevine monitoring is an important aspect of the viticulture industry, and remote sensing technologies are a powerful aid in reporting vegetation information for better vineyard management practices. However, the understanding of vine spectral responses as viewed by optical sensors has to be developed further, and was undertaken in this study.

Chlorophyll pigments drive photosynthesis, a biochemical process in plants, which contributes to physiological performance and productivity, making it an appropriate leaf characteristic for detailed examination. This study aimed to develop a thorough understanding of the relationship between (i) leaf-level spectral reflectance and transmittance properties and (ii) pigment concentrations, via ground-based sampling. This was achieved through the examination of two ground campaign tools, as well as current spectral data processing techniques and workflow methods. A spectrometer and SPAD chlorophyll meter collected non-destructive measurements during leaf senescence and grape harvest, and wet chemical extraction methods determined chlorophyll content (expressed in terms of unit leaf area and leaf fresh weight).

Reflectance indices, first order derivative indices, and a continuum removal approach were used to generate eighteen reflectance-based attributes. This study performed a series of chlorophyll estimation models through iterative ordinary least square regression, followed by two methods of model validation. Performance metrics indicated strong models with high explanatory power; the continuum removed depth normalized total area metric was presented as the optimal non-destructive attribute for accurate chlorophyll estimation for leaf level field

campaigns (R^2 = 0.93). Chlorophyll expressed in units of fresh weight yielded more consistent models than in units of leaf area. The chlorophyll meter also presented compelling results ($R^2 \geq 0.78$), and both sensors were determined to be appropriate for field validation campaigns for this vineyard study.

Table of Contents

List of Tables

List of Figures

Acknowledgments

As the heart of this thesis centres on wine, I think it only appropriate to raise a glass and propose a toast to the following:

I would like to acknowledge and thank the National Science and Engineering Research Council and the University of Victoria for their financial contributions to this effort. To the great people of Terra: for welcoming me, working with me, collaborating on this project, and their continued support over the years. To my other committee members, Mark Flaherty and David Goodenough: for their insight and guidance.

To Venturi Schulze: their open minds, willingness, and cooperation enabled this research to be conducted. Thank you Marilyn, Giordano, and Michelle for allowing me trample around your vineyard for a summer, cover it in flagging tape, and ask endless questions. It was a pleasure getting to know you, and an honour to watch your work ethic, perseverance, and dedication to your craft.

To my supervisor, and the top of the pyramid – Olaf. Thank you for creating and running an amazing lab, with such an eclectic group as us. It was incredibly rewarding and enjoyable working within HLRG, thank you for the experiences and for a home for the last decade.

To my family: my sister, for coming in at a crunch, and my mother, I could not have done this without you.

To my friends, far and wide: your support kept me sane, thank you for believing in me. Michelle Lai, Gina Martin, Laura Colquhoun, Celeste Dempster, Steph Blazey, Kylee Pawluk, Jessica Fritterer, Erin Langtham, Jason Garnham. David Schmidt, Erin Cusack & Greg Noel – for being in charge of fun. To Janet Sheppard, for listening.

Lastly, but close to my heart, to the comrades in the trenches with me: my lab mates. Rafael Loos, Fabio Visintini, Georgia Clyde, Roger Stephen. To the rest of the gang, both past and present. I'll finish with a special thanks to Geoffrey Quinn: for your tolerance of my endless stream of questions. Your guidance, teaching, and willingness to put up with me are greatly appreciated. It was a pleasure annoying you. Thanks buddy.

Dedication

I would like to dedicate this to my Mother, and to Joan Williams.

Chapter 1 Introduction

1.1 Context and Rationale

Remote sensing technologies provide a wide range of systems and methods to collect and process data for monitoring the environment (Jensen, 2007). These in turn are used to address a myriad of questions concerning vegetation identification (Schmidt & Skidmore, 2010), assessment (Niemann, Quinn, Stephen, Visintini, & Parton, 2015), quantification (Hall, Louis, & Lamb, 2003), and change detection (Desclée, Bogaert, & Defourny, 2006) at different ecosystem scales. In the past three decades, remote sensing technologies have added significant developments to viticulture management. Early efforts started with automated vineyard row identification and classification, and have progressed to complex health assessments and crop yield predictions (Zarco-Tejada et al., 2005). Remote sensing is a powerful aid as it remotely observes and records vegetation information, linking questions regarding better grape production to answers on physiological plant status.

Recently, advancements in precision agriculture have concentrated on improving localized crop management practices through custom treatments for differing areas of plant productivity, thus regulating crop variability, increasing performance, and maximizing yields (Cerovic et al., 2015; Uno, 2003; Zhang, Wang, & Wang, M. 2002). Additionally, localized treatments reduce consumption of water, pesticides, and fertilizers, but require a synoptic view of the entire property at a given time to determine where resources have to be allocated. Precision agriculture will be relied on more in the future to combat crop susceptibility from the increase in global and localized weather

pattern extremes (Johnson et al., 2012; Ustin, Roberts, Gamon, Asner, & Green, 2004). Remote sensing is one field that provides technologies and tools to better monitor these effects. Remote sensing data can be utilized to address three interrelated areas that are central to productivity models: (i) directly and easily communicates ancillary information to grape growers for better crop management, (ii) the pressure from governing bodies for better environment practices and policy implementations, (iii) the commercial aspect of increasing consistent yield.

The Canadian wine industry is young compared to its global counterparts, but is nonetheless prosperous. According to Statistics Canada, the sale of domestic wines has increased every year. It was estimated that 15.7 litres, or equivalent $220 worth, of wine were purchased per-capita last year (Statistics Canada, 2014). The latest reports estimate over 20 billion dollars in liquor sales countrywide, with roughly a third attributed to wine sales (Statistics Canada, 2015). Despite well-established vineyards in British Columbia and Ontario (Figure 1.1), the volume of imported wines continues to outpace domestic consumption, suggesting that perhaps production is not meeting consumer demand for domestic wines. However, as with any industry, it can develop only as quickly as the technology supporting it. Wine growers are dependent on innovative methods to monitor the wellbeing of an entire estate over the growing and harvesting season to produce the most desirable grapes.

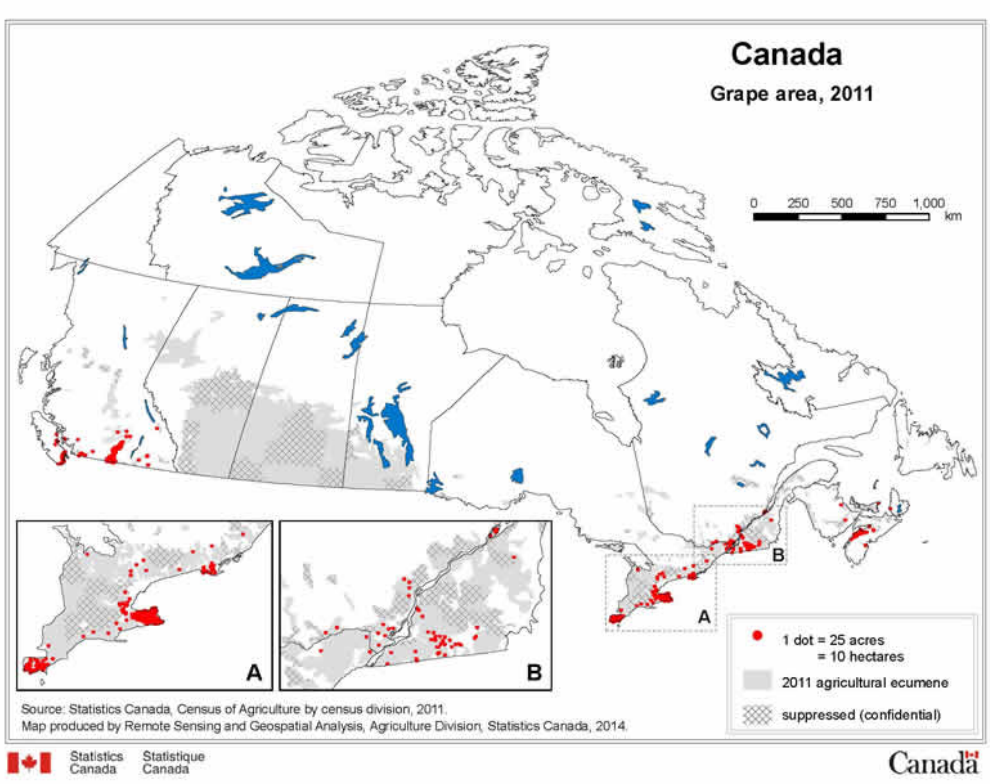

Figure 1.1: Distribution map of Canada's grape growing regions. British Columbia and Ontario exhibit highest density of viticulture plots. Adapted from "Thematic maps from the Census of Agriculture: Grape Area, 2011" Statistics Canada (2011). Copyright 2015. Reprinted with permission.

In British Columbia there are over 800 vineyards, with the vast majority located in the Okanagan region (Oliver, Osoyoos, and Kelowna) (British Columbia Wine Institute, 2012). In contrast to large, profit-driven agricultural crops, small-scale viticulture estates like those on Vancouver Island have adopted more ecologically mindful practices. This requires viticulturists to have extensive knowledge of the vineyard: vine growth, training, thinning, and an understanding of the *terrior*. Terroir is a French term specific to the wine industry, describing the subtle influences of environmental and human impacts on the wine taste profile (Robinson, 1994; Sommers, 2008). In *The Geography of Wine,* Sommers explains that each vintage is the culminated effort of viticulturists (grape growers) and vintners (wine makers), as well as vineyard location. The location

encompasses various environmental factors; climate, comprising of sunlight strength and duration, temperature, and rainfall; the soil, including chemical components, drainage, density and porosity considerations; and the topography, including slope gradations and hydrology. These variables contribute to the taste, bouquet, and colour of the wine. Balancing these factors, together with vine growth, maintenance, and management, demonstrates how labour-intensive the viticulture process is. The yield per unit area is not as large relative to other crops, but there is high value placed on the quality of wines.

With such a strong connection to these environmental elements, it follows that a sustainable, balanced relationship is a priority. Viticulturists are tasked with maintaining high crop quality and consistent yield, without pesticides or herbicides and with minimal use of fertilizers. Excessive use of pesticides and herbicides cause pest resistance in vines and residue in the wine (Lichtfouse et al., 2009). In addition, superfluous nutrient contamination leads to water pollution in surface and groundwater, inducing cultural eutrophication in water bodies (Daughtry, Walthall, Kim, de Colstoun, & McMurtrey, 2000; Saari et al., 2011; Stroppiana, Fava, Boschetti, & Brivio, 2011). Ongoing remote monitoring of crop development with airborne datasets benefits vineyard practices.

1.2 Study Scope

This study focuses on the assessment of plant productivity through field spectroscopy methods as they relate to foliar pigment content, and establishing the relationship between these two vegetation characteristics. While there are varying ways to define and quantify plant physiological status, biochemical estimation was the primary focus of this study. Pigments drive the principle biochemical process, photosynthesis, which ultimately determines plant productivity and physiological performance (Taskos et al.,

2014). There were two considerations in the study design to establish scope: the determination of a focal pigment for modeling, and the distinction of the remote estimation to be used, as it relates to the current body of literature in this field.

Chlorophyll *a* and *b* pigments are highly informative regarding vegetation productivity (Gitelson, Viña, Ciganda, Rundquist, & Arkebauer, 2005). Located within chloroplasts in the leaf palisade and spongy mesophyll, chlorophylls absorb incident radiation strongly in the blue (430-450nm) and red (660-680nm) regions of the electromagnetic spectrum (Cui, Vogelmann, & Smith, 1991; Zhang, 2011). These visible ranges provide optimal amounts of energy required for two photosynthetic processes: photosystem I and II (Campbell & Reece, 2002; Strever, 2012; Ustin et al., 2009). Pigment absorption features appear as depressions in the reflectance spectra (Jensen, 2005). Chlorophylls are the most abundant pigments, as there are five to ten times the concentrations of chlorophylls compared to other pigments in healthy vegetation (Almeida & Filho, 2004; Clevers et al., 2002; Kırca, Yemiş, & Özkan, 2006; Nascimento & Marenco, 2010; Xue & Yang, 2009). There are secondary pigments that include carotenoids; these yellow and orange pigments help dissipate high-energy incident light for the photosynthetic system. Additionally, anthocyanin (pink, purple, and red pigments) develops in some vegetation to protect against overheating and damage from UV light during leaf senescence (Sims & Gamon, 2002). These secondary pigments are responsible for only 25-35% of the absorption of incident radiation, while the chlorophylls account for more than 60-75% (Strever, 2012). Stress and senescence cause a more pronounced decline in chlorophyll than in other pigments, making it an optimal pigment for monitoring.

Nitrogen is meaningful to many agriculturists in the context of a fertilizer, and some studies choose to focus on it (Kokaly, 2001; Lamb et al., 2002; Li et al., 2010). However, nitrogen is not the primary factor in the photosynthesis process, but rather a secondary supporting biochemical (Taskos et al., 2014). In addition to the subtle but important distinction between the roles of chlorophyll and nitrogen, the intent of this study was to compare two instruments, one being a chlorophyll meter. This justified utilizing chlorophyll pigments as the concentration for this study; chlorophyll was used as the proxy for photosynthetic productivity within plants, as pigment content is a direct quantifiable measurement.

Remote vegetation physiological monitoring is used to establish plant health status, and is an important aspect of environmental management (Curran, Dungan, & Peterson, 2001; Schlemmer, Francis, Shananhan, & Schepers, 2005). Vegetation stresses and nutritional deficiencies are caused by both micro and macro environmental elements (Zarco-Tejada et al., 2005). Chlorophyll pigments are one foliar biochemical monitored as a proxy physiological indicator, as they are closely associated with the photosynthetic function within the plant tissue (Blackburn, 1999). This method of spatial variability tracking, and health assessment through pigment concentration determination is essential in aiding all crop development patterns (Haboudane, Tremblay, Miller, & Vigneault, 2008), and also specifically vineyard management (Hall, Lamb, Holzapfel, & Louis, 2002).

The passive collection of the electromagnetic spectrum in the visible and near infrared region is the primary non-destructive dataset used in this study. The utilization of optical remote sensing technologies captures a synoptic view of grapevine spectral

characteristics (Cunha, Marcal, & Silva, 2010; Fiorillo et al., 2012; Lamb, Weedon, & Bramley, 2004; Martín, Zarco-Tejada, Gonzalez, & Berjón, 2007). Hyperspectral datasets are collected as either individual samples (reflectance spectra) or as imagery (data cubes). Spectroscopy facilitates the examination of an object's composition without requiring its destructive determination (Vincent, 1997). The spectral reflectance of vegetation, both its shape and magnitude, is determined by the vegetation's biochemical and physiological composition (Gitelson, 2012). Pigments heavily influence the visible portion of the spectrum, while the near infrared (NIR) is influenced by water content and internal leaf cellular structure (Cho & Skidmore, 2006; Kokaly, Despain, Clark, & Livo, 2003). Reflectance measurements therefore are a non-destructive means to assess latent vegetation health variables through spectral responses dictated by plant productivity and physiological performance (Blackburn, 1999).

The study of biochemical estimation is fundamental in assaying vegetation stress and health, as pigments drive physiological function (Curran et al., 2001). For assessing remote estimation, this study investigates the varied chlorophyll content at a leaf-level scale within a Vancouver Island vineyard. If stress factors are unchecked, resource productivity is compromised. Factors such as heavy metal soil contamination, pest infestation, and lack of water contribute to stress, but are difficult to quantify and observe directly in vegetation spectral responses. However, chlorophyll is sensitive and well suited for measuring subtle productivity changes resulting from these stresses (Borengasser, Hungate, & Watkinds, 2008; Carter & Spiering, 2002). Chlorophyll is a quantifiable and appropriate measurement that acts for latent variables, such as plant productivity, stress, and plant aging (Kochubey & Kazantsev, 2007).

1.3 Research Motivations

The motivation for this study builds on the observation of Quinn (2010), stating that "ultimately, before the method can be determined an operational success or failure, an in-field validation must be conducted" (p.101). By extension, ground campaign data are critical for validation of airborne collections, as they are direct measurements to compare remote airborne imagery against, and to assess how accurately the imagery captures ground conditions. Airborne studies in turn facilitate better precision viticulture practices. Effective field data requires vetting of valid sampling techniques to ensure that measurements are appropriate for the study's objectives (McCoy, 2005). Laboratory measurements are one component of the field data collection. These traditional destructive methods of chlorophyll extraction are the keystone to linking non-destructive measurements to quantifiable results (Ruiz-Espinoza et al., 2010). The non-destructive measurements are the counterparts to the laboratory extractions, and consequently, the method of their processing is critical in capturing and characterizing the fundamental relationship between chlorophyll and spectral response. The three methods of (i) collection, (ii) extraction, and (iii) processing were the focus of this study. This proof-of-concept study evaluated the workability of current leaf-level field-based sampling techniques, field campaign tools, and processing methods for the remote estimation of grape leaf chlorophyll.

Some studies place greater emphasis on imagery correction and processing, leaving field methods ambiguous (McCoy, 2005). Field methods are problematic if not thoroughly conducted, resulting in error propagation. Milton, Schaepman, Anderson, Kneubühler, and Fox (2009) noted that the sampling strategy of a field campaign should be designed as a function of the end-use of the spectral measurements. Building on this,

there are three rationales for field measurements: (i) characterization of an individual object of interest, (ii) capturing the spatial influences on an object by sampling over a varying area, and (iii) calibration of an airborne dataset for atmospheric correction (McCoy, 2005). The original sampling design of this study includes all three aspects. Ground sampling measurements of different grape varietals were taken during the fall season, encompassing the majority of the vineyard, to capture spatial and phenological variations. In addition to sampling for full characterization of the grape leaves, field measurements were taken concurrently on the day of the airborne collection for validation.

1.4 Study Objectives

The purpose of this study is to conduct a ground based empirical study to develop a thorough understanding of the relationship between the leaf-level spectral reflectance and transmittance properties of vegetation, and pigment concentrations. This was achieved through the examination of ground campaign tools and current non-destructive data processing techniques. This study aims to address the following objectives by using pigment modeling to:

i) Develop an comprehensive understanding of the spectral responses occurring at leaf level and relate them to chlorophyll estimation

ii) Investigate current spectral and statistical processing approaches used in the literature, and apply appropriate methods to this study

iii) Determine which ground validation sensor is most appropriate for estimating chlorophyll in grapevines.

iv) Identify possible confounding and influential factors encountered within the processing streams

1.5 Processing Workflow

Two field campaign sensors, the Soil-Plant Analyses Development (SPAD) chlorophyll meter and the Analytical Spectral Device (ASD) FieldSpec Pro spectroradiometer, were used to collect in situ data at a leaf-level scale for *Vitis vinifera,* and converted to non-destructive representative metrics. Several different spectral methods of conversion were explored. The reflectance-based attributes, in addition to the chlorophyll meter readings, were evaluated as an estimate for chlorophyll content via regression modeling. In addition to the performance of the handheld sensors, the statistical processing streams of the sensor measurements were evaluated, and the workflow in Figure 1.2 was applied. The SPAD chlorophyll meter and the spectrometer were assessed and compared for their abilities to provide nondestructive estimates. Iterations of regression modeling established various causal relationships between remotely detected measurements and chlorophyll content. The optimal predictive model produced the least amount of model and residual error, and was validated using two common methods for additional performance evaluation.

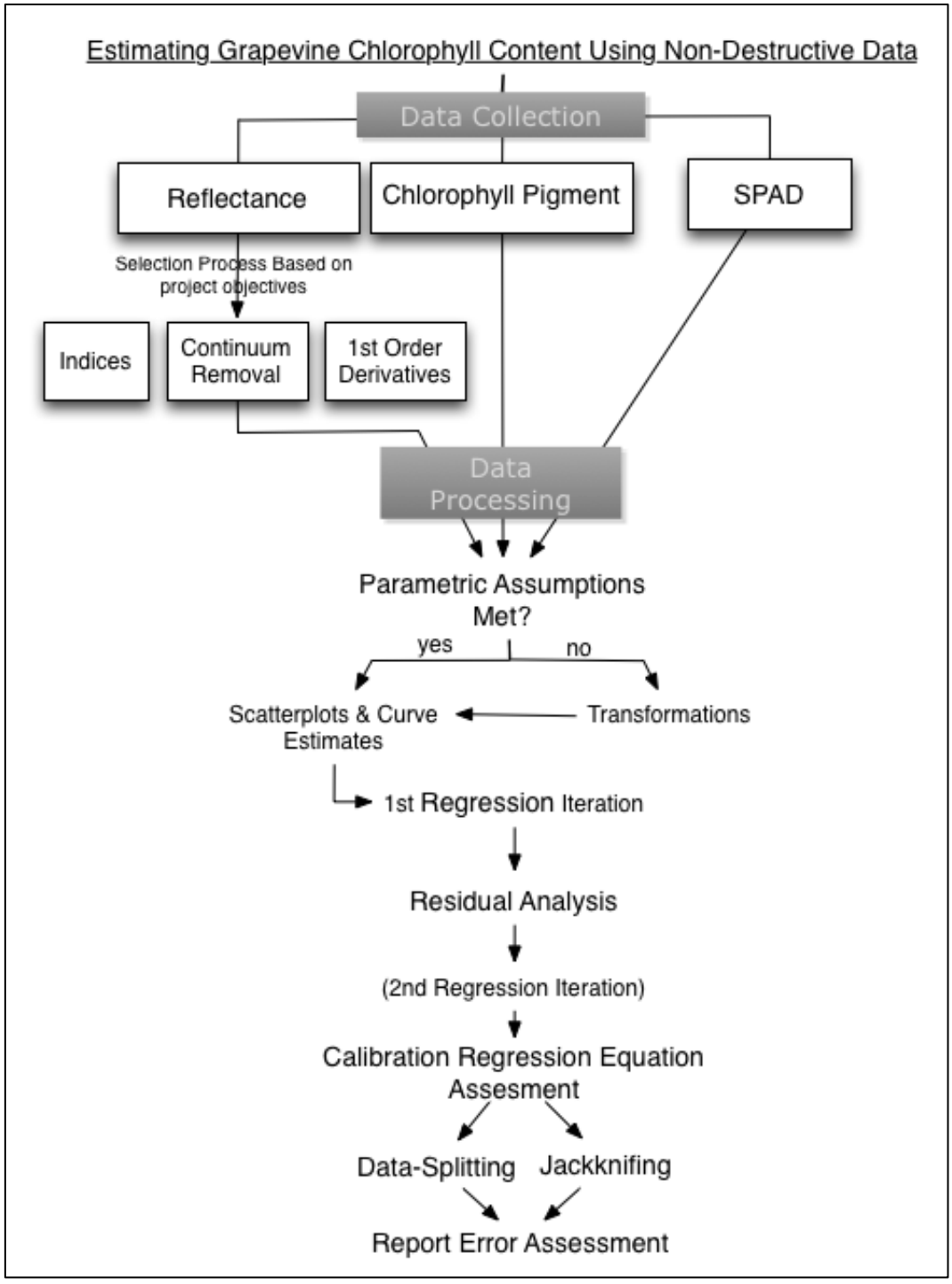

Figure 1.2: Processing workflow sequence for the non-destructive estimation of chlorophyll content as a plant productivity indicator using reflectance and chlorophyll meter measurements in conjunction with destructive pigment determination for regression modeling.

Chapter 2 Methods

2.1 Data Acquisition

This methodology chapter examines the descriptive statistics of the central datasets, as well as the pre-processing steps used in assumption verification prior to various parametric analyses. The statistical approach was an exercise in the application of current remote sensing technologies to explore the relationship between biochemical and non-destructive measurements, despite unforeseen sampling limitations that arose during the data collection.

2.1.1 Field Measurements

2.1.1.1 Sampling Site

The sampling site for this study was the Venturi-Schulze vineyard, located in the Cobble Hill region of the Cowichan Valley, on Vancouver Island, British Columbia. This part of the island is in the eastern rain shadow of the Vancouver Island Ranges, making the soil and mild temperatures well-suited for viticulture (Hynes, 2011). While the winters are less harsh than inland wine regions of British Columbia, so too are the summers less extreme in temperature ranges, reducing the number of growing degree days and requiring a longer ripening time (Danehower, 2010). The volcanic soils of the region are high in mineral content, and are known for bringing a distinctive flavour to the wines (Danehower, 2010; Hynes, 2011). A combination of water-retaining clay and limestone with drainage sands achieves optimal soil moisture for the vines. The vineyard operates on the principle of natural sustainability, the terroir being an integral part. To this end, there is a complete absence of herbicides or pesticides. Fertilizers are applied by hand, using organics such as a stinging nettle solution or kelp extract (Marilyn Venturi,

personal communication, May 2011). All vine pruning and berry harvesting is done manually. There is no irrigation system in place, thus the vineyard relies on rain and groundwater: a small gully borders the western edge of the estate and a small pond to the northeast (Figure 2.1). This requires the vines to grow long root systems to access the water table. This vineyard successfully balances the environmental conditions required to produce high quality grapes. The Venturi-Schulze vineyard is an example of an optimal, sustainable, agriculture model that encourages the natural elements to dictate management practices (Lichtfouse et al., 2009).

The Venturi-Schulze wine is comprised completely of estate-grown grapes, meaning there is no supplementation of grapes from other regions. The mixing of geographically different wines is a common practice among larger wineries to ensure large quantities of a consistent product through blending (Danehower, 2010). On Vancouver Island the grapes grown are typically more robust to cooler climates, such as the heartier German varieties. There are six white varietals, and three reds: Pinot Auxerrois, Pinot Gris, Kerner, Ortega, Medeleine Sylvaner, and Siegerrebe; Schönburger, Pinot Noir, and Zweigelt.

An independent vintage report by the BCWI (British Columbia Wine Institute, 2012) stated that 2011, the data collection year, was one of the coolest spring and summer on record, delaying bud break by two weeks and deferring the entire growing and ripening process. The danger with a late harvest is the increased risk of rainfall, causing the grapes to burst and for rot to set in, rendering the berries unusable for wine production (Robinson, 1994).

2.1.1.2 Sampling Strategy

The vineyard has over 8,000 vines on a 20-acre plot growing in a North-South direction. The property is on a western facing gentle slope with approximately 6 meters of elevation change. The estate is broken into two regions: the original vineyard consisting of the two western plots, and the larger portion of four plots developed later (Figure 2.1). The northern section of the original vineyard has the most limited light exposure, the sparest vine placements, as well as extensive mixing of varieties within rows. The other plots had more homogeneous groupings of varieties, and therefore were the focus of the study.

The field-based data was collected as part of a larger campaign, one which included the acquisition of airborne spectral and spatial data. The ground campaign was intended to provide ancillary data for the airborne dataset. It was originally envisioned

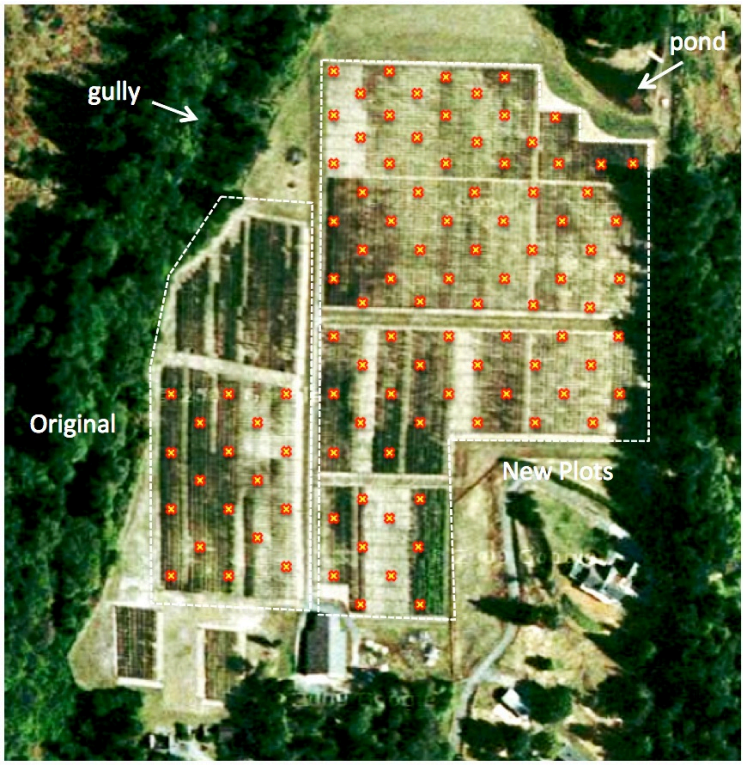

Figure 2.1: Aerial view of vineyard layout with approximate placement of repeat sampling sites for both SPAD and ASD measurements in red. Delineation of the original and new vineyard plots have been outlined, and naturally occurring water sources identified. (Google Inc., 2016)

that a ground-based, continuous-surface map be extrapolated from the chlorophyll point measurements to compare to the airborne imagery. As such, systematic sampling of the vineyard was selected over a random sampling design. While a truly random sampling avoids spatial bias, potentially it results in incomplete coverage or data clustering (Townend, 2002). Ultimately the airborne campaign was excised, and the ground-based sampling became the central focus, but the datasets were still viable with the altered project objectives.

Sample sites were established at discrete, equally distanced locations in a grid design that covered the spatial extent of the vineyard (Figure 2.1). Efforts were made to not oversample the edges of the plots. For consistency, all samples were measured from the western side of the vine. Leaves were selected near the top of the plant, from fruiting cane spurs. The top of the canopy was sampled in order to mirror the measurements captured by an airborne sensor. Because sampling was performed after veraison (the point in phenologic development when the grapes change from growing to ripening), lateral shoots off the spurs were included, as all major pruning was completed and no sample sites would be lost (Robinson, 1994). The individual leaves were identified with flagging tape for weekly, repeatable, in-situ measurements. Repeat collections spanned the month of September, and destructive sampling occurred in October. The number of sample sites had to balance complete spatial coverage with the time required to collect. As spatial coverage and not cultivars was the interest, the number of samples per variety ranged from six to fourteen.

After pooling all the varietals together, the number of sampling sites was determined sufficient for the statistical analyses of this study. Babyak (2004) presents a guideline of

at least 10 to 15 samples for every predictive variable in modeling. Alternatively, a base of at least 50 samples is advised, with the addition of a minimum of 8 more for every additional function variable. Notwithstanding the altered sampling design, the number of sample sites still fell within the recommended guidelines, especially as univariate analysis was used.

The two feasibility limitations for this project were time and accessibility. Of primary concern was the limited battery life of the instruments. Additionally, to compliment the airborne collection, the instruments were used within two hours on either side of the solar noon. This strategy minimizes the shadows in remotely sensed imagery. The second logistic was accessing the vines themselves. A green netting system was cast over the entire vineyard to deter birds and other wildlife from consuming the grapes as they ripen. The netting was situated on top of the large posts of the trellis system, and was buried around the entire perimeter of each plot. The netting was expensive and fragile, and required careful unearthing to allow the ASD spectrometer operator access to the individual vines. While moving down between the rows of vines was manageable with the equipment, moving laterally along the rows proved very challenging with an ASD system, as access was only possible between the large end posts and their anchoring wires.

In addition to the seventy-two leaves that were flagged and repeatedly measured, leaves were also collected at arbitrary points throughout the vineyard in order to supplement the data collection for pigment modeling. This ensured a complete and robust range of pigment measurements for the given collection period. The rationale was that when generating predictive models, the predictive powers are only as robust as the data

range, as model trends may change beyond the range of data collected. The grape harvest occurred over the month of October and was completed on November 1st. Then the flagged leaves were sampled destructively. Typically less than a day lapsed between grape harvest and leaf collection.

2.1.1.3 Spectroscopy Measurements

Readers may refer to the Analytical Spectral Devices Inc. manual (ASD Inc., 2002) for the complete specifics of the ASD spectrometer collection, but the following provides a summation. A fibre optic cable collected the light from the vegetation sample, where it was projected through a diffraction grating that breaks the light into individual wavelength components. Detectors converted the photon light energy into a voltage, and stored it in the computer as a 16-bit Digital Number (DN). The diffraction grating and the detector are collectively known as the spectrometer. There are two types of detectors within this system. The first collects in the 350-1050nm range (the visible and near infrared). This detector had a fixed photodiode array, where individual wavelengths have a corresponding location along the array. The spectral resolution was 3nm, and determined as the Full Width Half Maximum (FWHM). Conversely, the 1050-2500nm region (the short-wave infrared), was collected with oscillating diffracting gratings. The detectors collected data in sequential measurements rather than simultaneously, as was the case in the fixed array. Two oscillating gratings had a 150nm overlap to ensure no data loss: the first spanned the 900-1850nm region; the second the 1700-2500nm region. The spectral resolution of these two regions had a wider Gaussian dispersion, between 10 and 12nm.

The ASD instrument was optimized prior to sample collection, allowing gains, offsets, and integration times to be detected automatically for the sensor. The surrounding light environment determined the sensor collection time to ensure adequate signal strength without saturation. Optimization utilized a white reference panel (or reference standard), which endeavored to emulate perfect spectral diffusion. The sensor itself generated a dark current, a type of systematic noise created by the electrical current of the machine. It was subtracted from the total signal of each channel (wavelength component). The instrument closed a shutter and collected the dark current, (i.e. the photons generated with no active vegetation measurement). The dark current was collected whenever a white reference measurement was taken.

Following the collection of the white reference and dark current DN's, the white reference was used for calibration. The ratio of the white reference against itself produced a baseline across the whole spectrum (value of 1). The software, RS^3, then generated subsequent relative reflectance values of the vegetation expressed as a ratio of the light from the vegetation and the light from the white reference. This was the final measurement that was collected and stored by the computer.

To collect a vegetation sample with the spectrometer, a contact probe with a leaf clip attachment was connected to the fibre optic cable; this provided a controlled environment to isolate the spectral response exclusively to the leaf and minimizes ambient light. The white reference panel was located on one side of the swiveling head of the leaf clip, and once the reference was collected, the leaf clip pivoted to a black backing. The leaf sample was placed against the contact probe, and the clip closed around the sample. The contact probe had its own light source. The adaxial surface of the leaf was always sampled.

Primary and lateral veins were avoided, as the vascular bundles have different constituents than the lamina area (Strever, 2012). Random noise can be reduced through averaging; 30 measurements were averaged to generate one sample, for 10 iterations. Post processing averaged the 10 samples into a single reading. Thus, each sample consisted of an average of 300 measurements.

2.1.1.4 Chlorophyll Meter

The Soil-Plant Analyses Development (SPAD) is a hand-held instrument that measures the transmittance of light through leaves (Steele, Gitelson, & Rundquist, 2008b). The SPAD was an alternative to reflectance-based sensors (such as the spectrometer) for estimating chlorophyll content. Once a relationship was established between the SPAD measurements and the destructive chlorophyll content of the leaf sample, direct non-destructive leaf chlorophyll estimates were possible.

The chlorophyll meter operates under the principles of the Beer-Lambert Law. Individually, Lambert's law stipulates that two identical media of equal sampling thickness will absorb an equal fraction of the light energy traversing it, all other factors being equal (Biochrom Ltd., n.d.-a). In other words, the proportion of light absorbed by a medium is independent of the incident intensity of that light. For example, for a medium that transmits 50% of the incident light energy: light starting at 100% strength (i.e. the light source) and proceeding sequentially through four discrete but identical media, will drop from 100% to 50% to 25% to 12.5%. Beer's Law further adds that the degree to which the light is absorbed is directly proportional to the concentration and path length of the medium through which it passes (Fadock, 2011). Beer's law determines the fixed

proportion value of the previous law (given as 50% in the example). Combining the two principles results in the equation of absorbance, *a*, expressed as

$$a = \log_{10} I_o/I, \quad (1)$$

where I_o is the intensity of the incident light and I is the intensity of the transmitted light (Biochrom Ltd., n.d.-a). This combined Beer-Lambert Law addresses the exponential attenuation of light through a medium (Curran, Dungan, Macler, & Plummer, 1991; Nascimento & Marenco, 2010). However, this principle is based on an 'ideal optical system' (a homogeneous medium), and internal leaf structures are vastly more complex, with air to liquid boundaries and inconsistent distribution of chlorophyll absorbers (Markwell, Osterman, & Mitchell, 1995; Uddling, Gelang-Alfredsson, Piikki, & Pleijel, 2007).

The sensor measures absorbance at two wavelengths, 650nm and 940nm to determine the SPAD output (M). The former wavelength location corresponds to chlorophyll absorption within the leaves, and the latter acts as a reference wavelength uninfluenced by pigments and compensates for varying leaf thickness (Netto, Campostrini, Oliveira, & Bressansmith, 2005; Richardson, Duigan, & Berlyn, 2002; Spectrum Technologies Inc., 2009). The equation to determine SPAD values is given as

$$M = k\log_{10} (I_{o650}*I_{940})/(I_{650}*I_{o940}), \quad (2)$$

where I_o is the incident energy at the given wavelength, and I is the transmitted light energy at the given wavelength (Cerovic, Masdoumier, Ghozlen, & Latouche, 2012; Uddling et al., 2007). The moving arm of the sensor closes around the vegetation sample, nestling it between the detector(receiving window) and light emitter (emitting window), avoiding ambient light interference (Figure 2.2). The incident measurements were

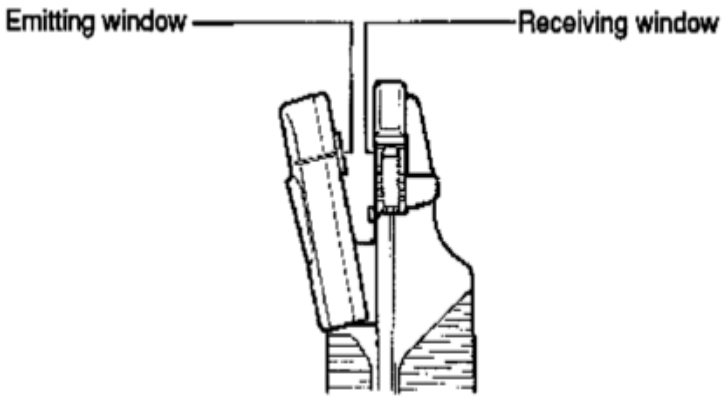

Figure 2.2: The emitting and collection arms of the SPAD chlorophyll meter provided in user manual (Spectrum Technologies Inc., 2009)

collected first without a vegetation sample. The receptor sensor within the receiving window stored the voltage measured from the red (640nm) and infrared (940nm) photons emitted from two light emitting diodes (LEDs) housed in the emitting arm (Spectrum Technologies Inc., 2009). The vegetation sample was placed then in the sampling slot and the voltage recorded from the photons transmitted through the vegetation sample. The sensor reported the unitless SPAD value as determined from equation (2). The accuracy of the sensor is reported at ± 1 SPAD value, with a repeatability of ± 0.3 SPAD value (Ruiz-Espinoza et al., 2010; Spectrum Technologies Inc., 2009).

2.1.2 Laboratory Measurements

The leaves sampled with the spectrometer and SPAD were sealed in labeled freezer bags, and placed in a cooler with ice. The samples were transported and processed in the university laboratory generally within a five-hour window of being collected. Reflectance spectroscopy and SPAD measurements were sampled in the exact manner as the field measurements, the only difference being they were performed in a darkroom laboratory to ensure limited and consistent environmental conditions.

2.1.2.1 Pigment Extraction

The extraction protocols for pigments were critical in gaining an understanding of the behaviour of pigments under different environmental conditions (Torres et al., 2014). Chlorophyll pigments were processed spectrophotometrically in solution (Wellburn, 1994). Once the non-destructive laboratory measurements were completed, circular leaf cores were removed using a cork borer until 0.0100g (±0.001g) of leaf material was obtained. A Mettler Toledo AL204 analytical balance was used to weigh the samples, and was reported to have a measurement repeatability of ±0.0001g (Mettler-Toledo, 2012). Sampling was within the same area as the non-destructive measurements to ensure the foliar chemistry extracted was as complementary to the non-destructive measurements as possible. Leaf material was transferred to a labeled test tube and kept in darkness to limit the degradation of the photosensitive samples. An additional dozen leaf disks were cored and weighed for moisture determination, and placed in plastic test tubes (empty weight also recorded). These samples were dried in a food dehydrator for two days, then sealed with screw caps and reweighed. The test tube weight was subtracted from the final weight to determine the leaf moisture.

Leaf core area measurements were determined for the destructive samples intended for both pigment and moisture determination. Sampled were scanned and WinSEEDLE software (Regent Instruments Inc., Canada) was used to determine the leaf area, and presented two methods of determination (Figure 2.3). The method selected was a more accurate representation of leaf total area, as it allowed user control and verification through photo comparison. There were up to ~1cm^2 difference between the automated and user confirmed measurements (Table 2.1).

Table 2.1: Comparison of two area measurements reported with the Winseedle software. Top 8 largest discrepancies reported, Sample 56 being the largest. Automated processes generally overstated vegetative areas

Sample ID	Sum of Automated Areas (mm^2)	Sum of Areas (mm^2)	Difference
Sample 56	734.4419	636.6145	**97.8274**
Sample 144	485.6959	410.0186	**75.6773**
Sample 108	618.4355	543.9495	**74.486**
Sample 54	418.0800	361.3155	**56.7645**
Sample 31	704.7575	650.6004	**54.1571**
Sample 76	677.2115	625.0254	**52.1861**
Sample 92	611.9605	569.6003	**42.3602**
Sample 165	506.2399	466.13	**40.1099**

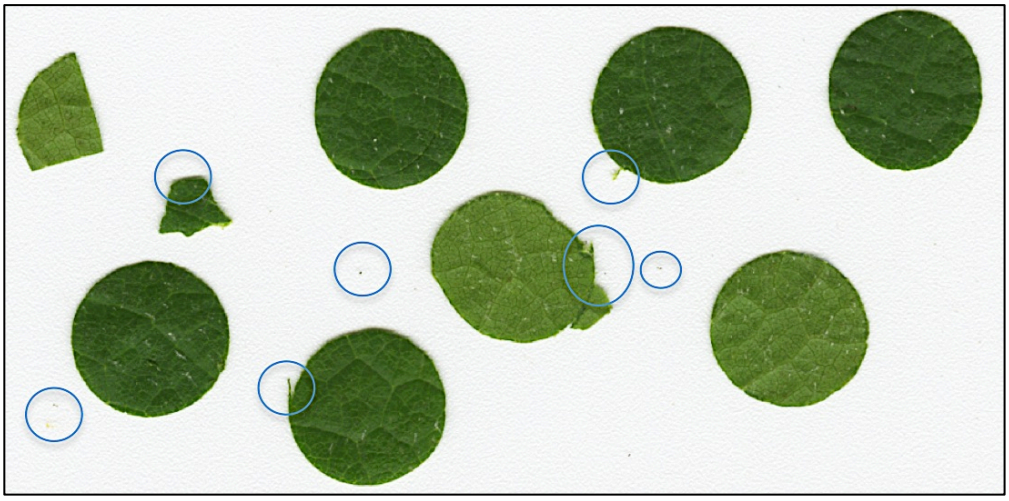

Figure 2.3: Images captured by scanner and Winseedle software determined area measurements. Images were used to validate the number of individual areas that the software identified as vegetation. Sample 56 demonstrates areas of shadow, textured edges, and small debris (blue circles), which were included in the automated summation of the total area.

The samples were transferred to the chemical laboratory facility. Within the confines of a fume hood, an Eppendorf Research-Plus adjustable volume pipette was used to dispense 10mL of dimethyl sulfoxide (DMSO) into each test tube. For every

batch of samples processed, a sample containing only pure DMSO was included as a control. Samples were placed in a water bath at a constant temperature of 60-65 °C for eight hours to facilitate extraction of the pigments.

There are numerous organic solvents that can extract pigments, and DMSO was determined to be the most appropriate for this study. DMSO does not require maceration of the leaf tissue, has slower sample degradation, and depending on plant material, extracts equally or more completely than other agents (e.g. ethanol, acetone, or dimethylformamide) (Hiscox & Israelstam, 1979; Sumanta, Haque, Nishika, & Suprakash, 2014). Minocha et al., (2009) noted a brown discolouration to the solvent, indicating material other than chlorophyll (such as other biochemical or non-pigment cellular matter) was extracted in the process. However, this discolouration was not observed in the present study. Only one study investigated a pigment extraction method specifically for grape leaves; however, this study by Lashbrooke, Young, Strever, Stander, and Vivier (2010) used a high performance liquid chromatography (HPLC) profiling method, and did not use DMSO as a solvent.

After the allotted incubation time, leaf disk material was removed prior to transferring 5ml of the suspension with the pipette into plastic cuvettes. Disposable pipette tips were used for single suspension transfers to ensure no cross contamination of samples. A Biochrom WPA Lightwave II UV/Visible spectrophotometer with a bandwidth of 5nm collected the pigment absorption measurements (Biochrom Ltd., n.d.-b). A blank sample comprising of pure, heat-treated DMSO acted as a reference to calibrate the sensor, having a complete absence of plant matter. Two sets of measurements were recorded and averaged from the spectrophotometer. The first was the absorbance at five user-specified

wavelengths (480, 645, 649, 663, and 665nm). The second was an absorbance wave scan, wherein an absorption spectrum was collected for a user-defined range (400 to 950nm).

There are several prominent extraction equations within the literature for determining pigment content. Equation suitability is contingent on the pigment absorption maximum and minimum, the solvent used to extract the pigments, the vegetation type, and the spectral resolution of the spectrophotometer (Wellburn, 1994). Given these specifications, the equations of Wellburn (1994) were employed in this study.

$$Chl_a = 12.19A_{665} - 3.45A_{649} \quad (3)$$

$$Chl_b = 21.99A_{649} - 5.32A_{665} \quad (4)$$

$$C_{x+c} = (1000A_{480} - 2.14(Chl_a) - 70.26(Chl_b))/220. \quad (5)$$

Chl_a is the chlorophyll *a* concentration, Chl_b is the chlorophyll *b* concentration, and C_{x+c} is the total carotenoid concentration. In addition to Wellburn's equation, absorption at 645nm and 663nm were recorded to apply to Arnon's (1949) alternative equation. Gao (2006) demonstrated Arnon's and Wellburn's equations produced almost identical results for different mediums (90% acetone vs DMSO) when considering the chlorophyll absorption feature from 600-700nm. Both equations reported pigments in terms of suspension concentrations (μg/ml), and were converted to leaf pigment content by dividing by either the fresh leaf weight (μg/g) or leaf area (μg/cm^2).

2.2 Data Pre-processing

2.2.1 Leaf Spectral Behaviour & Data Reduction

To confidently work with spectral reflectance data, an understanding of the influences on, and behaviour of, the spectral response is required. The relationship between spectral reflectance and pigment abundance is inversely proportional, and attributed to the strong

absorption of energy by pigments (Tucker, 1979). The visible portion of the electromagnetic spectrum spans 400nm to 700nm, and encompasses the photosynthetically active region (Steele, Gitelson, & Rundquist, 2008a; Yoder & Pettigrew-Crosby, 1995). Spectral sensitivity to this chemical energy production system provides important information on vegetation status (Sims & Gamon, 2002). This region is governed by the quantity and dispersion of various foliar pigments. The red range, 630-690nm, is influenced heavily by the absorption of chlorophyll *a* and *b*. The blue range, 400-500nm, is influenced by both chlorophylls and carotenoids (for example, carotene and lutein), as both pigment groups have local maxima absorptions in this region (Lichtenthaler, Gitelson, & Lang, 1996). Energy in the green spectral region is not absorbed effectively by chlorophyll, and thus is reflected back, instead of utilized by the photosynthesis system (Campbell & Reece, 2002).

Leaf optical properties change at different phenological stages, corresponding to pigment abundance. For example, as vegetation senescences the chlorophyll content decreases, as does the strength of the green reflectance, allowing the reflectance in the orange and red regions from carotenoid pigments to dominate leaf colour. The most prominent feature in vegetation spectra is the Red Edge (RE). It is characterized by the change in spectral magnitude from strong pigment absorption in the red region to the high spectral response in the near infrared (NIR) region. The latter is influenced by leaf thickness, water content, and light scattering rather than by pigment content (Strever, 2012). Plants evolved to scatter and reflect NIR energy, as it does not supply sufficient energy to overcome the threshold requirement to run photosynthesis and generate water and sugars (Atwell, Kriedemann, & Turnbull, 1999). Internal and surface structure of

leaves, as well as wall-to-air cell boundaries, contribute to the degree of scattering (Datt, 1999). Beyond the RE, water content is a physiological feature also affecting spectral response, and is influential in the following regions with energy absorption features: 970 nm, 1190 nm, 1450 nm, 1940 nm, and 2500 nm (Sims & Gamon, 2003). At larger scaled investigations, such as the canopy level, there are additional factors influencing the spectral response. These include changing canopy bidirectional reflectance-distribution functions (BRDF), spectral mixing, atmosphere interference, sensor and illumination angles, and canopy complexity to name a few (Cochrane, 2000). While these factors are recognized as contributing components of spectral imagery, the focus of this study will be kept at the leaf-level scale.

Complete characterization of spectral features is possible due to the high spectral resolution of hyperspectral data. Consequently, there is high data volume and redundancy, as a band at any given location provides similar information to its adjacent neighbouring waveband (Thenkabail, Enclona, Ashton, & Van Der Meer, 2004). Therefore, approaches have been developed with the intent of reducing dataset redundancies while still retaining the significant components of the spectral responses (as outlined in the previous paragraphs), without a loss of critical information. This study presents three methods of dataset reduction: reflectance indices, 1st order derivative indices, and a continuum removal approach. Additional alternative methods of data manipulation are available; one example is Principle Component Analysis (PCA), and involves the projection of data to new coordinates that maximizes variance between uncorrelated variables (Thenkabail, Lyon, & Huete, 2011). These data mining algorithms

are excellent tools for descriptive modeling, but they are very dataset-specific, making inferential modeling more of a challenge.

2.2.2 Indices

2.2.2.1 Index Categories

An index is a mathematical manipulation of data to convey a single unitless metric that represents a key data characteristic (Jago, Cutler, & Curran, 1999). Numerous variations of indices have been prototyped and modified for spectral processing to depict vegetation biophysical and biochemical characteristics. Indices have evolved over decades in remote sensing research, and can be assigned to general categorical classes. These include, but are not limited to, the following:

- Single Band (SB): This index utilizes only one band to represent a spectral feature. This value could be the reflectance or derivative, or a mathematical manipulation of them (e.g., the inverse of the reflectance, or the difference between two bands).
- Simple Ratio: At its most simplistic, this category is the ratio of two bands. In modified forms, the numerator and/or denominator may be the difference, summation, or product of two bands. Alternatively, a constant is sometimes added as an offset or adjustment factor. Simple ratios are often used to normalize a particular spectral characteristic.
- Normalized Difference (ND): This ratio consists of a two-band difference in the numerator and a band summation in the denominator. There are at least two band combinations in these functions. Modifications can include the addition or subtraction of additional constants to correct or account for specific foliar characteristics.

- Three Band Indices: The objective of these indices is to avoid excess influence of single bands through mathematical combinations of three bands and adjustment constants. These include both modern hybrids of ratio and orthogonal indices, as well as developments of traditional ratio indices to include a third spectral region for better biochemical characterization.
- Other Mathematical Manipulations: This category comprises the range of mathematical determinants that do not specifically fall into the above categories. Examples include the determination of areas under spectral curves, summation of data ranges, and differences without the use of ratios, or the combination of two or more of the above methods in combination.

These general classes were the result of a broad literature review, but were influenced particularly by (Broge & Leblanc, 2000; Girma, Skidmore, de Bie, Bongers, & Schlerf, 2013; Haboudane, Miller, Pattey, Zarco-Tejada, & Strachan, 2004; Jackson & Huete, 1991).

In addition to the groups described above, there are further approaches to categorize indices. Depending on the research focus, indices can be grouped and used according to the scale for which they were developed and intended. Many of the compounding effects subject to airborne spectral imagery are controlled at leaf-level, especially studies utilizing contact probes. A subset of Vegetation Indices (VI) intended for canopy scale have developed soil-adjustment methods to account for and correct these compounding factors. While some indices are general enough for across-scale application, these VIs are tailored to spectral image processing, as they require additional information apart from

the vegetation spectra (i.e. soil spectra). Indices of these variations were not considered in this study.

Bandwidth is an additional way indices can be categorized. Broadband spectra were standard in previous development of indices, and continue to be applied today. However, due to their widths, some multispectral datasets do not have the spectral resolution required for modern hyperspectral-derived indices. Within the context of this thesis, hyperspectral indices are considered to be a metric that capitalizes on the high spectral resolution. For example, the Continuum Removed Total Area (CRTA) and first derivative Red Edge Inflection Point (REIP) metrics are hyperspectral indices that fall under continuum and slope transformation categories exclusive to hyperspectral processing. With that distinction, indices presented in this section are considered multispectral, while hyperspectral methods are investigated in subsequent sections.

2.2.2.2 Index Selection

Criteria were established to refine a selection process for indices to be included in this study. With numerous proxy measurements for vegetation physiological status, such as stress or health estimates, there are subsequently many indices devoted to their determination. They were not developed for the direct purpose of chlorophyll estimation however, and therefore, did not meet the study objectives. Furthermore, only unaltered reflectance measurements were used in indices for this section; index metrics from derived spectra, such as first order derivative, are investigated in subsequent sections.

The literature was reviewed for robust studies that identified indices that consistently performed well. The selection process design outlined the following criteria: (i) Studies had to investigate chlorophyll directly. Ozelkan, Karaman, Candar, Coskun, and Ormeci

(2015) undertook an extensive index review concerning grapevines, but considered ratios of CO_2 and leaf area as a photosynthesis proxy, rather than chlorophyll pigments. (ii) Airborne studies required a ground campaign component with a comparable chlorophyll estimate method. Therefore studies like those of Haboudane et al. (2008) which focused on canopy Leaf Area Index (LAI) metric estimates, were not included, as well as studies that looked solely at large scale data simulations, such as PROSPECT and SAIL modeling (Jacquemoud et al., 2009). (iii) To balance both recent advancements as well as the historical multispectral indices, studies were limited to the last two decades. (iv) Studies had to compare indices rather than demonstrate the abilities of a newly proposed index. A minimum comparison of ten indices was required, and consequently works such as Gitelson, Gritz, and Merzlyak (2003) or Zhang, Chen, Miller, and Noland (2008), while significant, were not included. While the process was not stringent, it aimed to identify indices in each of the category groupings, without redundancy of any one type, and identify well performing robust indices.

These requirements yielded five key studies: Haboudane et al. (2008); Le Maire, François, and Dufrêne (2004); Lin, Popescu, Huang, Chang, and Wen (2015); Main et al. (2011); and Xue and Yang (2009). These studies met all the criteria, but nonetheless were very different from each other in their objectives and concentrations. Lin et al. (2015) highlighted the differences in chlorophyll estimation between healthy and water stressed leaves, while Main et al. (2011) conditioned all regression models to be linear for a panoptic comparison of estimation metrics. Le Maire et al. (2004) compared a simulated dataset with a field campaign, and Haboudane et al. (2008) investigated the application of indices specifically to agricultural crops (corn and wheat). The indices appearing in two

or more studies as a superior performing index are included in Table 2.2. Indices that did not consistently perform well within all papers, but yielded the best results within their respective study are presented in Table 2.3.

Table 2.2: Indices common to more than two of the studies that met the present study's index selection requirements.

Name	Equation	Reference
Datt	$(R_{850} - R_{710})/(R_{850} - R_{680})$	(Bisun Datt, 1999)
Maccioni	$(R_{780} - R_{710})/(R_{780} - R_{680})$	(Maccioni, Agati, & Mazzinghi, 2001)
mND_{705}	$(R_{750} - R_{705})/(R_{750} + R_{705} - 2R_{445})$	Sims & Gamon (2002)

Table 2.3: Indices that did not overlap within all five studies, but were the top performing within their respective referenced paper. Performance was reported with coefficients of determination and RMSE in most cases.

Name	Equation	Source Reference
ECI1	$(R_{645} - R_{455})/R_{REP}$	Lin et al. (2015)
Vogelmann	R_{740}/R_{720}	Vogelmann et al. (1993)
OSAVI	$(1 + 0.16)*(R_{800}-R_{670})/(R_{800} + R_{670} +.16)$	Rondeaux et al. (1996)
MTCI	$(R_{750} - R_{705})/(R_{709}-R_{681})$	Dash and Curran (2004)
$\frac{TCARI}{OSAVI}$	$\frac{3 * ((R_{700} - R_{670}) - 0.2 * (R_{700} - R_{550}) * (R_{700}/R_{670}))}{\frac{(1 + 0.16)*(R800-R670)}{(R800 + R670 +.16)}}$	Haboudane et al. (2002)
Gitelson2	$(R_{750}- R_{800}/R_{695} - R_{740})$ -1	Gitelson et al. (2003)
mSR_{705}	$(R_{750} - R_{445})/(R_{705} - R_{445})$	Sims & Gamon (2002)

The indices selected for analysis were Datt, Maccioni, modified NDVI, and Gitelson2. Priority was given to Gitelson2 from the second tier of potential indices (Table 2.3), as it was developed specifically with grape leaves by Steele (2007), and was anticipated to perform well in the present study.

2.2.3. Derivatives

The first order derivative is simply the slope of the reflectance spectrum for a given wavelength range (Curran et al., 1991; Mutanga, Skidmore, & Prins, 2004). The most prominent feature is the steep change in reflectance values between spectral bands, located at the abrupt change from pigment absorption in the visible to the structural scattering in the NIR (Cho & Skidmore, 2006). Derivative datasets provide the location, magnitude, and shift of the maximum slope, known as the red edge inflection position (REIP). Utilization of the REIP is an appropriate method for hyperspectral data, as the shift is often within ±25nm, and difficult to capture with coarser multispectral sensors (Quinn, 2010; Ustin et al., 2004). The link between the REIP and chlorophyll is well established from previous laboratory studies (Gitelson, Merzlyak, & Lichtenthaler, 1996; Lichtenthaler et al., 1996). Derivatives are reported to be less sensitive to illumination changes (Tsai & Philpot, 1998), and therefore preferred to reflectance, due to its insensitivity to variability from the surrounding environment (Blackburn, 2007). That being said, this study used a contact probe and leaf clip on the spectrometer, minimizing irradiant light and controlling the fixed light source. It was of interest to see if the derivative could produce superior estimates compared to reflectance, when these illumination issues were not of paramount concern.

Prior to determining the slope, the spectra underwent smoothing to reduce random spectral noise (Smith, Steven, & Colls, 2004). Calculating the derivative was a balance of spectral feature discernment and preservation, with data noise reduction through smoothing. Tsai and Philpot (1998) presented three methods of smoothing: Savitzky–Golay smoothing, Kawata–Minami smoothing, and mean-filter smoothing. The moving kernel filter was appropriate for this dataset, as the random noise was minimized previously through sample averaging.

Derivative values were determined by a finite approximation equation (Tsai & Philpot, 1998; see also Huang, Turner, Dury, Wallis, & Foley, 2004). Alternative methods included high-order curve fitting, which imposed either a third-order polynomial or a Gaussian distribution curve on the reflectance profiles to estimate expected derivative values (Baranoski & Rokne, 2005; Kochubey & Kazantsev, 2007; Lamb et al., 2002). However, these methods were developed for coarser spectral resolution data. The slope was therefore approximated by the use of a five-band moving average kernel using the equation

$$D\lambda_i = (R\lambda_{i+2} - R\lambda_{i-2})/(\lambda_{i+2} - \lambda_{i-2}), \quad (6)$$

where $D\lambda i$ was the first derivative at wavelength i, and $R\lambda_{i+2}$ and $R\lambda_{i-2}$ were reflectance values two bands adjacent to i on either side. λ_{i+2} and λ_{i-2} were wavelength values incrementally adjacent to i by two intervals on either side. The visible region was significantly more varied than the red edge region, and therefore a five-band kernel was chosen for the additional smoothing it provided in without causing any spectral feature loss.

2.2.4 Continuum Removal

Spectral feature analysis methods often focus on the magnitude, location, or shift of local absorption maxima and minimum. Continuum removal examines the spectral shape of a feature in its entirety (van der Meer, 2004). The continuum is the 'background absorption' onto which other absorption features are superimposed (Clark, 1999). Clark and Roush (1984) identified the continuum as a straight line fitted to the local maximum points, or peaks, of the spectrum (Huang et al., 2004). Figure 2.4 illustrates this concept, as the convex hull above the spectrum can be thought of as the reflectance that would occur when no absorption is present (Huang et al., 2004).

Originally developed for geological identification and classification purposes, Kokaly and Clark (1999) adapted the continuum removal approach for laboratory-based vegetation spectrometry, although chlorophyll was not initially included. Curran et al.

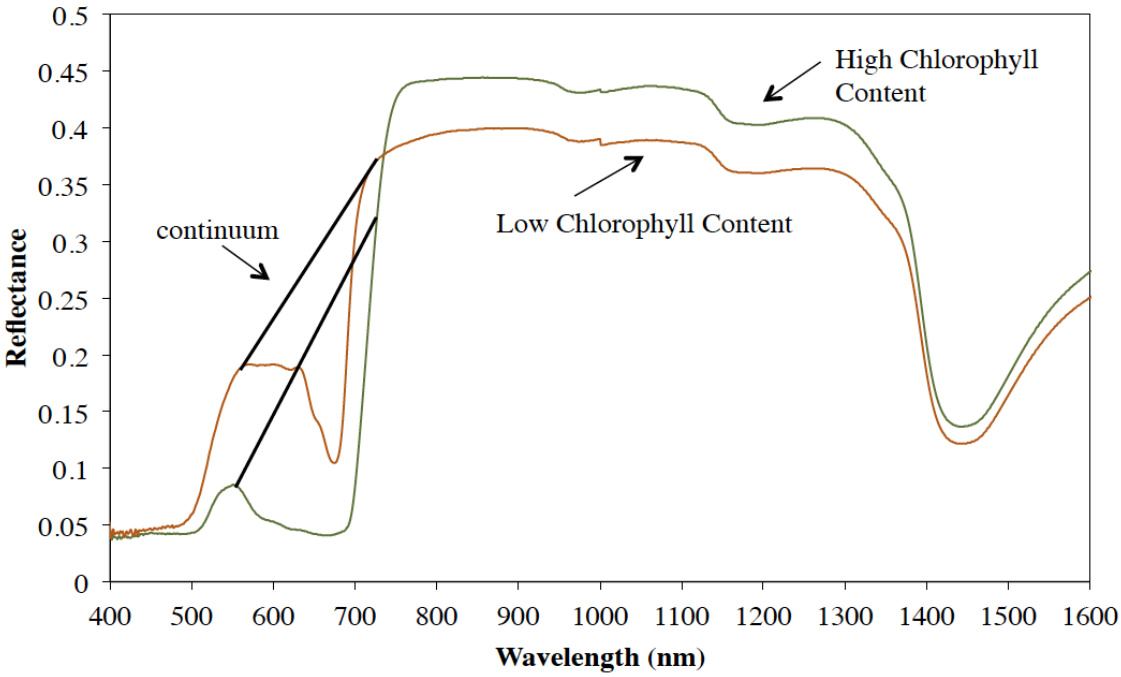

Figure 2.4: Samples of varying spectral reflectance with chlorophyll content of grapevine leaves. Continuum lines have been added for the red absorption feature

(2001) expanded the study to include photosynthetic pigments, namely chlorophyll *a*, and *b*, and total chlorophyll. Both of these studies were conducted with dry vegetation processing, and developed to include fresh canopies by Mutanga, Skidmore, Kumar, and Ferwerda (2005).

While a number of papers employ continuum removal methods, a portion of which cover vegetation applications, a markedly small body of work specifically investigates chlorophyll estimation with continuum-removed reflectance. More often, studies apply continuum removal for vegetation health and stress, quality, or density (Broge & Leblanc, 2000), focus on other biochemical constituents or water features as their primary vegetation indicator. The subtle difference between chlorophyll and stress estimation is demonstrated in Sanches, Filho, and Kokaly (2014), who developed the plant stress detection index (PSDI). This index targeted continuum removed values at the centre of the chlorophyll absorption feature and on the green peak edge, which relate to the present study both in processing methods, and in spectral areas of interest. Despite these similarities, attempts to use this index to model chlorophyll did not yield strong results, as the index was intended to identify stressed plants, not pigment content estimation. Continuum removal metrics such as these were not included in the present study.

The continuum-removed spectrum is determined as the ratio of the reflectance to the continuum line at each individual wavelength. Band depths generated at the centre of the absorption feature, as well the area under the band depth curve, are two frequently used metrics (Curran et al., 2001; Kokaly & Clark, 1999). Data is normalized to allow for analysis and comparison across datasets; this is achieved by dividing each continuum-

removed spectra by the maximum band depth ratio value of the region, allowing all values to be scaled between zero and one. The specifics of the processes to generate the metrics used in this study are elaborated on in section 2.4.3.3.

2.3 Statistical Overview

2.3.1 Parametric Considerations

This section addressed statistical issues surrounding parametric analysis; namely, data distributions and applied methods to correct them. As the purpose of the study was to investigate the appropriateness of non-destructive data measurements as representatives of plant productivity, the assumptions of parametric regression were the focus.

Two assumptions of parametric analysis were confirmed: (i) the measured samples were ratio data, and are therefore considered continuous; (ii) the dependent variables, extracted chlorophyll measurements, were independent of each other as they were destructively sampled. Functionally, it is understood that a leaf's spectral response is dictated by foliar pigment content; however, for the purpose of non-destructive estimation modeling, pigments are treated as the dependent variable and the reflectance and transmittance measurements as the independent variables.

There are varying opinions about which assumptions take precedent, and which can be violated with minimal influence on the statistical test of interest (Field, 2009). One notion concerns the importance of normal frequency distributions for variables prior to running regression models. In cases where the distribution of model variables is considered less important (Rawlings, Pantula, Sastry, & Dickey, 1998), priority is given alternatively to variance stabilization of model residuals. However valid this assumption, many disregard it with the rationale that as long as the residual dispersion is normal, the distribution of the input variables matters little (Bates & Watts, 1988). This study attempts to normalize

frequency distributions for both the independent and dependent variables in regression analysis, as well as the model residuals.

Kolmogorov-Smirnov (K-S) and Shapiro-Wilk (S-W) tests were used to assess the frequency distribution of the variables. While there are additional tests to examine normality (e.g. Anderson-Darling), these two compare the dataset to a simulated normal distribution frequency. The simulated data represents the population of the original dataset and has the same mean and standard deviation. It also assesses any significant differences in distributions (Field, 2009). The S-W was the more powerful test when considering varying sample sizes and possible distributions (Razali & Wah, 2011). All tests were conducted at a 0.95 confidence ($p = 0.05$), to determine whether the null hypothesis was rejected or accepted. The null hypothesis (H_0) and the alternative (H_1) for the K-S and S-W test are (Townend, 2002):

H_0: no significant difference was determined between the distributions of the populations that the sample data came from and that of normally distributed simulated data

H_1: a significant difference was determined between the distributions of the populations that the sample data came from and that of normally distributed simulated data.

The non-destructive metrics would pass the K-S test but fail the S-W test on occasion. The primary difference between the tests is that the Shapiro-Wilk is more stringent and sensitive to deviations from the expected distribution, especially in scenarios involving kurtosis or bimodal distributions, which were reoccurring within the study (Rawlings et al., 1998). The K-S test is less stringent in such situations, especially when considering larger sample sizes. When the test results differed, K-S results were given precedent, as

there are no data manipulations to correct for bi- or multi-modelism. To minimize skewness and correct for frequency distributions, transformations were applied.

Data transformations, such as log or square root transformations, correct for data distribution complications by altering the relative relationship between variable samples. This modifies the data range and distribution between samples without affecting the sampling order (Field, 2009). Data transformations also address anomalous values, allowing these samples to be representative of the population without their removal from the dataset (Townend, 2002). Transformations are applied in two instances: (i) as a convenient means to adjust skewness in a frequency distribution, and (ii) to linearize a curvilinear relationship between two variables. The latter avoids adding regression model variables that do not contribute to its predictive power. Transforming the data to obtain normal distributions was preferred to non-parametric analysis, as it was not as likely to discover subtle differences between metric performances and fully characterize the relationships between the dependent and independent variables (Townend, 2002). Only transformations intended to correct a single curvature in the line estimate (one-bend transformations) were attempted (Rawlings et al., 1998). The $\log_{10}$, natural log (ln), square root and inverse were used predominantly in this study.

The final parametric consideration was the identification and exclusion of outliers, or points that were inconsistent with the rest of the dependent or independent observations (Rawlings et al., 1998). Outliers have the potential to influence and leverage regression models, biasing the mean and inflate the standard distribution (Field, 2009). While all samples contribute to the model development, ideally any sample removed from the model would not change the model estimate parameters drastically. All datasets

were assessed for outliers; if a sample was identified as an outlier in one dataset (e.g. laboratory reflectance measurement), the sample was removed from all datasets (SPAD and pigment extraction measurements). This avoided data gaps and any possible error propagation. Initially chlorophyll datasets were normalized and checked for samples that fell beyond two standard deviations, considered the limit for outliers (Field, 2009). The sample dataset was standardized using the equation

$$x = \frac{x_i - \mu}{\sigma}, \tag{7}$$

where each sample's z-score was determined by subtracting the sample (x_i) from the mean (μ) and divide by the standard deviation (σ) (Field, 2009). The samples that fell farther than two standard deviations from the mean were shortlisted for spectral examination to determine if they were suitable for inclusion in the study. Samples that completely lacked a chlorophyll absorption feature, and rendered reflectance metrics unusable, were excluded from the study. As they were so far senesced, instead of contributing to modeling, they severely restricted the regression capabilities. A total of five samples were excluded from the study, leaving 134 samples for further analysis.

2.3.2 Statistical Checks

The data descriptive statistics assessed the suitability of each sampled dataset for parametric modeling. Initial adjustments to the data were made and unanticipated characteristics, such as data skew, were accounted for. All statistical analyses were performed using a combination of SPSS v.20 (IBM Corp., 2011) and Aabel v.3 (Gigawiz Ltd., 2008) statistical software.

2.3.2.1 Chlorophyll Pigments

In this study, chlorophyll measurements were reported in micrograms of pigment content per unit of fresh leaf weight (μg/g) and in micrograms of pigment content per unit of fresh leaf area (μg/cm^2). Many studies report one or the other, preference depending on the basis of the study (i.e. an agricultural or a spectroscopy focus). No consensus was reached on which was more appropriate.

2.3.2.1.1 Chlorophyll by Fresh Weight

The laboratory analytical chlorophyll extraction returned a range of 1.48 μg/mL to 17.36 μg/mL for chlorophyll pigments in the suspended solution for the grape leaves. This equates to an approximate range of 150 -1750 μg/g chlorophyll pigments per unit of fresh weight leaf matter. When comparing data from other studies that spanned varied conditions, species, and collection methods (Table 2.4), the data sampling fell within an acceptable range of chlorophyll content.

In an ideal normal distribution, both skewness and kurtosis would be zero, indicating symmetry within the data spread. For the chlorophyll pigments by fresh weight, skewness and kurtosis were reported as +0.281 and -0.858, respectively. This slightly positive skew

Table 2.4: Summary of chlorophyll content by unit fresh weight from a selection of studies to demonstrate expected pigment ranges.

Study	Chlorophyll Range μg/g	Vegetation Type
Ruiz-Espinoza et al. (2010)	170 - 700	Basil leaves
Quinn (2010)	133.5 – 1299.0	Douglas fir needles
Uddling et al. (2007)	0 - 600	Wheat Birch and Potato
Yoder and Pettigrew-Crosby (1995)	200 - 1600	Big Leaf Maple
Curran et al. (2001)	56 - 1760	Slash Pine needles
Strever (2012)	560 - 1760	Grapevine leaves

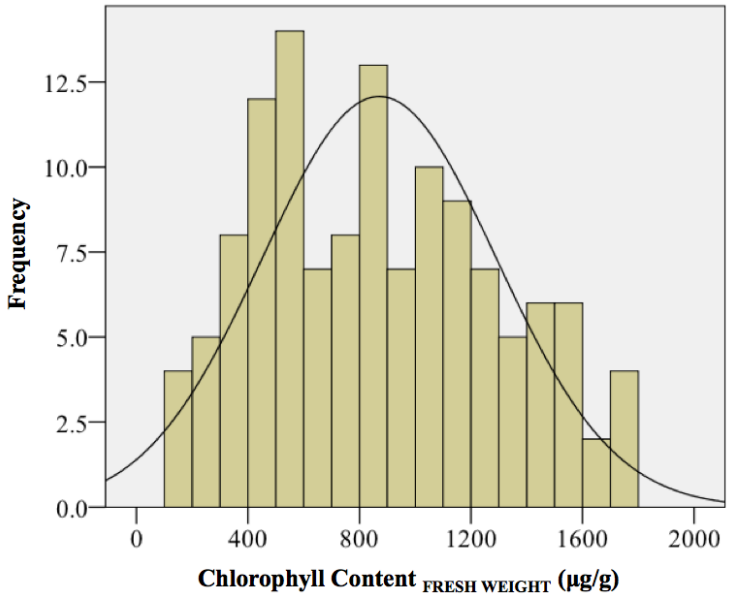

Figure 2.5: Histogram of chlorophyll content by fresh weight frequency distribution, indicating a slight positive skew and extended tail on the right side of the distribution

indicated a tendency for a small percentage of the data to extend further into the higher values (Figure 2.5), which caused an elongated distribution tail on the right.

The skewness in the distribution was explained by the data collection occurring over senescence; there were higher occurrences of leaves with lower chlorophyll content. Strever (2012) noted a sharp decline in chlorophyll content in the late stages of grape leaf phenology. The data for this study was collected at the end of that pigment decline, which resulted in the majority of the samples having reduced chlorophyll content, but with a small number not yet showing signs of senescence and continuing to exhibit higher chlorophyll values. A negative kurtosis suggested a flatter distribution than the ideal normal distribution. However, the histogram revealed it was not distinctly platykurtic. The data appeared to be multimodal, due in part to the binning of X-axis values. Regardless, the modes indicated a lack of results in two particular value ranges, rather than in multiple fundamentally different sampling populations.

The K-S and S-W tests reported a rejection of the null hypothesis ($p = 0.01$ and 0.00), both considerably below the alpha α 0.05 significance level, which indicated a deviation from normality. To correct the distribution, a square root transformation was used, with new data values passing the K-S test ($p > 0.05$).

2.3.2.1.2 Chlorophyll by Area

The laboratory analytical chlorophyll extraction returned a range of 1.48 μg/mL to 17.36 μg/mL for chlorophyll pigments in the suspended solution. This equated to an approximate range of 2.0 – 36 μg/cm^2 chlorophyll pigments per unit area of fresh leaf matter. Given the range from a selection of other studies (Table 2.5), the data sampling fell below the expected upper range of values. This result was not anticipated, as the pigment content per unit fresh weight matched the published literature. Xue and Yang (2009) reported a chlorophyll range less than the present study, while four were at least 14 μg/cm^2 higher in their upper extents.

Four studies (Ruiz-Espinoza et al., 2010; Strever, 2012; Uddling et al., 2007; Xue & Yang, 2009) expressed chlorophyll pigment in terms of both unit area and unit fresh

Table 2.5: Summary of chlorophyll content by unit leaf area from a selection of studies to demonstrate expected ranges.

Study	Chlorophyll Range μg/cm^2	Vegetation Type
Ruiz-Espinoza et al. (2010)*	9.80 – 51.0	Basil leaves
Uddling et al. (2007)*	0.00 – 80.0	Wheat, Birch and Potato
Xue and Yang (2009)*	2.30 – 24.5	Leafy green crop vegetables
Steele (2007)	0.30 – 51.0	Grapevine leaves
Strever (2012)*	0.00 – 50.0	Grapevine leaves

* indicates studies that had chlorophyll expressed in both unit area and unit fresh weight

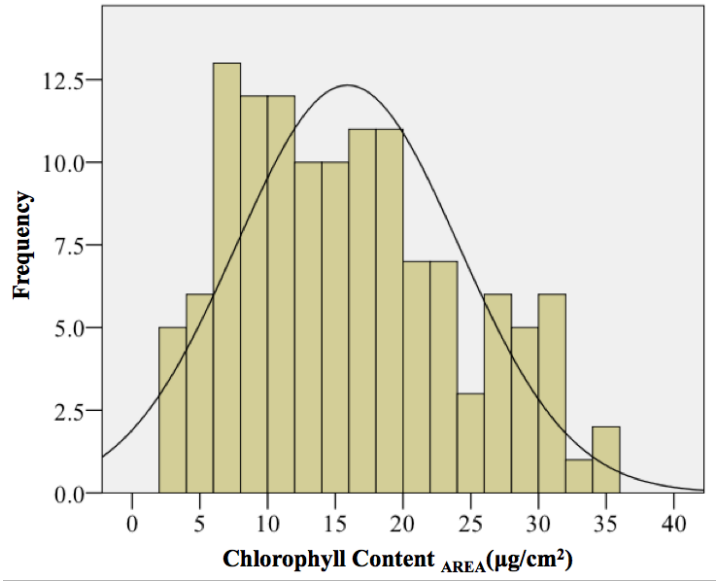

Figure 2.6: Histogram frequency distribution of chlorophyll content by fresh leaf area, indicating a slight positive skew and extended tail to the right, and a bimodal distribution.

weight, and while the present study compared well in terms of the fresh weight range, the chlorophyll by unit area differed.

The mean, skewness, and kurtosis were reported as 15.89 µg/cm^2, +0.448, and -0.659 for the chlorophyll pigments by unit area. The positive skewness resulted in an extended frequency tail into the higher pigment concentration values, as was the case when chlorophyll was expressed in terms of fresh weight. The area measurements did not appear as multi-model as their fresh weight counterparts, as there was only one prominent separation in the dataset (~ 25 µg/cm^2 in Figure 2.6). A S-W test for normality confirmed a significant deviation from normal ($p = 0.02$), however, the more lenient K-S test did not ($p = 0.09$). The rationale behind the distribution was the same as the fresh weight measurements, as the relationship was intrinsically the same. To correct the distribution, a square root transformation was applied to the pigment values, which resulted in the

frequency distribution achieving normal distribution in both tests ($p = 0.20$ and 0.15 for K-S and S-W, respectively).

2.3.2.1.3 Pigment Measurements Comparison

The most fundamental measurement of chlorophyll for this study was as a concentration; the quantity of pigments within a given amount of suspension (the solvent being DMSO). The inclusion of fresh weight or area gave the measurement context to vegetation and remotely sensed data by expressing it in relatable terms. However, it introduced an error of uncertainty accumulated through the weight and area determination. Nonetheless, the latter is one of the most commonly reported units of measurement, as leaf area is linked to the vegetation surface from which reflectance measurements are taken.

As anticipated, the relationship between the area and fresh weight foliar measurements was strongly positive (Figure 2.7). However, the data were heteroscedastic; there was increased dispersion of points along an axis in relation to the other variable, indicating a non-constant variance (Hayes & Cai, 2007). There was a clear increase in sample dispersion as chlorophyll content increased, resulting in an escalation in the difference between the two measurements at higher chlorophyll concentrations. It was originally hypothesized that the area measurements introduced error through compounding error accumulation in leaf disk measurements that could be responsible for the difference between chlorophyll measurements. However, the heteroscedasticity did not correspond to the samples that had larger area measurement discrepancies. No correlation between those two factors was found. The samples with the highest residual values deviating from zero were 62, 58, and 74, and corresponded with samples that

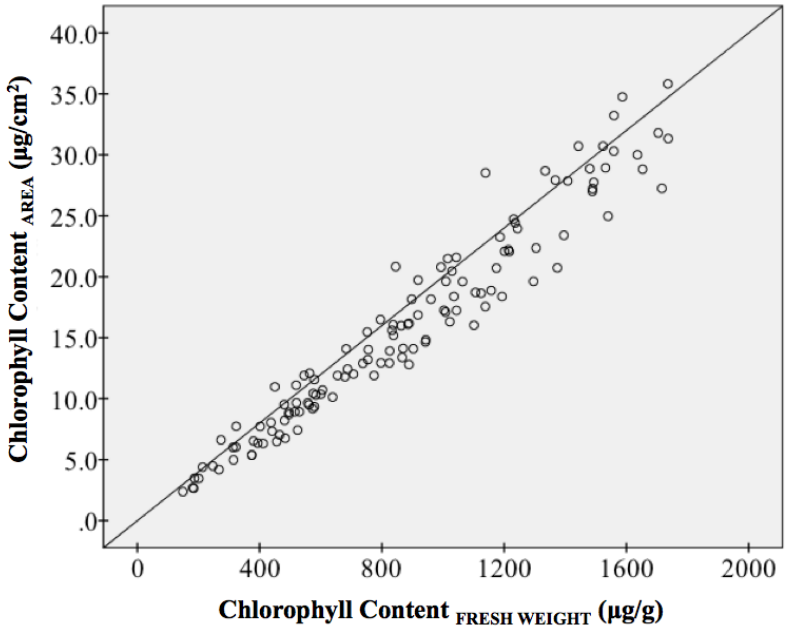

Figure 2.7: A scatterplot visualizing the relationship between two methods of expressing chlorophyll content: by per unit fresh leaf area or by per unit of fresh leaf weight. Samples varied more at the higher chlorophyll contents

deviated furthest from the 1:1 ratio line of Figure 2.7. Comparatively, Samples 56, 108, and 144 had the highest area discrepancies (Table 2.1). Therefore, errors in the area measurement were not the largest contributing factor for the difference between area and fresh weight measurements. Further explanations were explored in section 4.2.

2.3.2.2 SPAD

In contrast to the extracted pigments, the chlorophyll meter dataset behaved as anticipated. The SPAD measurements were considered suitable for parametric modeling based on the two test for normality ($p = 0.20$ for K-S test, and $p = 0.77$ for the S-W test). The data range, 7.24 to 52.6 SPAD units, was compared to similar studies (Richardson et al., 2002; Steele, 2007; Uddling et al., 2007). The mean, skewness, and kurtosis were reported as 28.26, +0.133, and -0.141. A histogram of the SPAD frequency distribution values indicated the spread was unimodal.

2.4 Data Analysis

2.4.1 Regression Considerations

2.4.1.1 Model Variables & Workflow

Non-destructive SPAD and reflectance measurements were used in the remote estimation of chlorophyll within grapevine leaves. All regression modeling utilized the extracted chlorophyll pigment content as the dependent variable; however, it is recognized that the foliar biochemical is functionally the explanatory variable. Pigments dictate how the spectral response will behave, and is technically an independent variable. In order to establish and build predictive relationships, destructive measurements had to be taken, and the pigments were treated as the dependent variable to be remotely and non-destructively determined. The purpose was to ascertain the amount of variation in the dependent variable was explained by the independent metrics, and whether that amount was significant. The hypothesis for all regression models presented in this paper were based on the following null and alternative:

> **H_0**: the independent variable (non-destructive metric) does not sufficiently explain the variation in the dependent variable (chlorophyll pigment content)
>
> **H_{alt}**: the independent variable (non-destructive metric) explains a significant portion of the variation in the dependent variable (chlorophyll pigment content).

Descriptive modeling characterizes the relationship between variables, while inferential models are intended for potential estimation of values other than those used in its development, and as such, undergo a validation process (Mutanga et al., 2004). The predictive models in this study were developed as candidates for future research. Furthermore, the statistical checks of the 'Line of Best Fit' approach required an iterative

process, as diagnostic testing often led to improvements to the model (Rawlings et al., 1998). This process is summarized in the bottom half of the workflow in Figure 1.2.

This study used single variate analysis, although multivariate analysis (partial least squares, multivariate, and stepwise) are alternatives used within the literature (Mutanga et al., 2005; Schlerf et al., 2003). However, those approaches require consideration as the results are more sample dependent and hyperspectral data processing is particularly susceptible to overfitting. Earlier studies commonly used stepwise linear regression for pigment estimation (Curran et al., 2001; Grossman et al., 1996; Kokaly & Clark, 1999). This method is less prevalent now as it was established that the iterative search for predictors took advantage of type I error rates (Field, 2009).

2.4.1.2 Reporting with Performance Metrics

In order to distinguish optimal models, comparison of performance metrics is necessary (Fadock, 2011). The coefficient of determination, R^2**,** reflects overall function of the explained to unexplained proportion of data (Field, 2009). The adjusted R^2 was reported, as this shrinkage technique is a type of validation that incorporates the model's potential for a new sample set (Babyak, 2004). The standard error of the model estimate, or the root mean squared error (RMSE), was also included. While the SPSS software provided the RMSE in linear regressions, the nonlinear equivalent was determined by the equation

$$\sqrt{SS_r/n}, \tag{8}$$

where SS_r is the sum of squares of the residuals and n is the number of samples.

In addition to these performance metrics, the graphical scatterplots of the residuals determined the appropriateness and success of a model (Rawlings et al., 1998). Residuals were standardized and plotted against the dependent variable to assess variance change

along the dependent range. The residuals mean value was confirmed to equate to zero (indicating no model over- or under-estimation), and K-S and S-W tests were used to confirm a normal distribution (equal, uniform spread around the predictive line estimate). Residual patterns can indicate outliers, unaccounted variables, or more complex relationships that are still present in the data after model fitting (McKillup, 2012; Ranjit, 2006). When the dispersion indicated nonconstant variance, a Breush-Pagan SPSS syntax macro statistical test was performed to determine heteroscedasticity.

2.4.1.3 Equation Predictor Parameters

When considering univariate analysis, the distinction between linear and nonlinear regression must be highlighted. Linear regression is often preferred because of the mathematical simplicity of its determination. Rawlings et al. (1998) asserted that, when all things being equal, "the simplest model that describes the observed behaviour of the system should be adopted" (p.398). There is some reticence regarding curvilinear modeling, and attempts to establish a linear relationship when it is inappropriate for the dataset can be limiting (Field, 2009). Caution regarding polynomial functions centres on the idea that the predictive powers of an equation are artificially inflated with the addition of equation variables (x, x^2, x^3, etc.) but contribute little to the explanatory power of the model. Overfitting refers to instances where more emphasis is placed on anomalies of the dataset rather than the underlying relationship between independent and dependent variables, and this bias is undesirable (Babyak, 2004). This was of particular concern as the explanatory power of the model was, ultimately, its most critical aspect concerning this study. Residual interpretation was pivotal in the application of both linear and curvilinear functions.

One of the limitations when undertaking large index assessment review studies is the batch processing approach to statistical analysis that is often required (Le Maire et al., 2004; Main et al., 2011). Batch processing ensures the same method is applied to all samples so that the results are comparable and do not contain bias. Yet not all studies that use this approach account for whether the true relationship between the two variables is different from the proposed functions, or if the models suffer from overfitting due to too many non-explanatory model variables. In the work done by Le Maire et al. (2004), every regression was a third order polynomial. Similarly, in work by Main et al. (2011), all of the regressions were linear. Employing this approach favours relationships that match the batch function. While both of these studies provide invaluable information as an overview and for general discussion, more consideration must be given to the nature of variable relationships. A specific model function was determined independently for each of the metrics in this study.

2.4.1.4 Model Validation

Significant attention was paid to model tuning throughout the study, and a validation process was applied to assess how well the models would perform with a different dataset. Given the scope of the study, an additional external validation dataset was infeasible; therefore, an internal validation was used (Harrell, 2015; Rawlings et al., 1998). The dataset was divided into a calibration group that would be used towards the model development, and a test group for model validation. An additional statistical permutation method, jackknifing, was also performed to assess the model's prediction error (Efron & Tibshirani, 1993).

As the SPAD generated only one unitless value measurements, a single regression model was performed for chlorophyll estimation. Conversely, reflectance measurements were reduced through continuum removal, derivatives, and indices to generate a suite of metrics for regression modeling. The best reflectance metric and the SPAD metric were put through a rigorous statistical assessment. In order to evaluate the coefficients obtained from regression modeling, as well as the effects of individual observations on the model, two techniques assessed the effectiveness of the regression equation at predicting chlorophyll estimates.

There are numerous ways of determining the optimal ratio to split the datasets into calibration and validation, but this depends largely on the objective of the model and the dataset size. Le Maire et al. (2004) split their experimental database in half (53:52 samples), as did Steele (2007) for his grapevine study (49:44 samples). In three studies by Mutanga et al. (2003, 2004, 2005), a 75-25% split of the data (72:24 samples) was used for calibration and validation sets. Williams (2008) suggested partitioning data into three independent sets with a 40:30:30 or 70:15:15 split for training, validating, and testing. While three groupings are possible for data mining methods with large datasets, a separate testing dataset was not a practical option for this study. Instead, the calibration dataset was used to test different model parameter settings and variable suitability, and then assessed on the validation set by determining the unbiased error and performance metrics.

A universal split was applied to all datasets collected in this study to separate the calibration and validation datasets. If a sample was segregated to the training dataset, it was done so for the SPAD, reflectance, and chlorophyll pigment measurements. The data

were stratified by chlorophyll concentration before data-splitting. While a random partitioning is the only way to ensure an unbiased process (Williams, 2008), this could critically limit the model's ability to fully characterize the range of chlorophyll. As this characterization was a vital component of this study, and as chlorophyll is nonlinear in nature with respect to reflectance, full range in both calibration and validation datasets were required to ensure optimal models (Gitelson et al., 2003). Even and odd sampling numbers to segregate half the data into training and testing datasets, a method replicated from Steele (2007). The ratio of training to testing was based on the rationale that with over 130 cases, the relationship could be well characterized with half of the data, leaving an equal amount for the assessment of the modeling, so that neither were favoured with more samples.

The second validation method was a resampling approach that compared the test statistic from a data subset to the test statistics of a reference distribution created from the data themselves, rather than a standard statistical distribution (Legendre & Legendre, 1998). An advantage of this method, aside from its computational robustness, is that it does not assume a specific distribution (van Belle, 2002; Vogt, 1993). The reference distribution was obtained by randomly permuting the data, without any consideration to whether the data was representative of a larger statistical population. Because of the nature of resampling, it was limited to observing bias and standard errors, and making inferences regarding only that specific dataset (Legendre & Legendre, 1998).

Jackknifing and bootstrapping are resampling methods designed for smaller datasets, where each test iteration consists of a subsample of the original data used to generate a test statistic, which is then compared to statistics for the entire dataset. The key difference

between bootstrapping and jackknifing is that the former resamples from the original observations with replacement (Carsey & Harden, 2014). This enables the resampling to approximate the original dataset as closely as possible, and to increase robustness and illuminate possible sampling bias. Jackknifing is a more systematic method that does not resample observations that have already been excluded in the subsample iterations (Legendre & Legendre, 1998). For this study, jackknifing resampling was performed, as it was representative of which samples were responsible for deviating from the line of best fit, and thus causing a weaker relationship.

2.4.2 SPAD

There were two objectives for the chlorophyll meter component of this study. The first developed two predictive models for chlorophyll estimation, using SPAD values and their corresponding destructive chlorophyll measurements. These models captured the difference between chlorophyll expressed in units of both leaf fresh weight and per unit leaf area. Ruiz-Espinoza et al. (2010) investigated SPAD and pigment relationships for basil leaves and found a higher coefficient of determination when pigments were expressed per unit area rather than in unit fresh weight. Uddling et al. (2007) found varying results for birch, wheat, and potato, depending on how chlorophyll was expressed, but chlorophyll expressed in terms of fresh weight had a stronger relationship when all species were considered together. Results will contribute to the support or contradiction of these findings. In the second objective, SPAD estimation models presented in the literature were assessed with the current datasets. These evaluated the chlorophyll meter, its suitability as a universal agricultural tool, and how the current grapevine dataset compared to those from the literature. Two published regression

models determined chlorophyll content on an area basis, as the vast majority of studies SPAD presented chlorophyll measurements in this fashion.

While the majority of studies utilized the SPAD as a validation tool in the scope of a larger scaled study (Baldy et al., 1996; Fiorillo et al., 2012; Lausch et al., 2013; Wu, Niu, Tang, & Huang, 2008), some investigated the abilities of the chlorophyll meter expressly (Marenco, Antezana-Vera, & Nascimento, 2009; Markwell et al., 1995; Monje & Bugbee, 1992). Lin et al. (2015) argued that because the SPAD sensor collects transmittance rather than reflectance measurements, it would be more sensitive than reflectance to water deficits. They demonstrated that models overestimated chlorophyll estimation in water stressed leaves of the camphor tree. However, Taskos et al. (2014) reported contradictory results, stating that within an irrigated and non-irrigated vineyard, the SPAD readings (and corresponding chlorophyll measurements) were unaffected by water application. This suggested that either more extreme water stress was required to observe deviations, or alternatively, plant species and structure affected SPAD readings more than moisture. Moreover, that finding lends weight to the argument by Markwell et al. (1995) that SPAD modeling be species specific for optimal prediction. Adding to discrepancies found between studies, no firm consensus has been reached on whether the relationship to chlorophyll measurements is linear or nonlinear. Both were reported, with nonlinear models characterized as exponential, polynomial or rational functions (Azia & Stewart, 2001; Broge & Mortensen, 2002; Lin et al., 2015). Linear models are considered superior to nonlinear models, as then a larger range of chlorophyll could be estimated. With curvilinear functions, SPAD values would not continue to increase at the same rate

as the chlorophyll content, suggesting a limitation in the sensor's ability to discern chlorophyll content as the sensor's measurements saturate.

SPAD studies from the last two decades were compiled, and equations for converting SPAD values to chlorophyll content were selected, based on the objectives, suitability, and success of the model. SPAD values were entered into the published models, and the resulting estimated chlorophyll was linearly regressed with the measured chlorophyll, to assess the strength and accuracy of the model estimation. Details of published models are elaborated in section 3.1.4.

2.4.3 Reflectance

2.4.3.1 Reflectance Indices

For decades, developments in non-destructive proximal measurements advanced optical indices for chlorophyll discernment at different geographic scales for vegetation monitoring (Malenovský et al., 2006). The indices below were identified as suitable for the estimation of chlorophyll for the present grapevine study. They covered most of the index class categories presented in section 2.2.2.1: The mND_{705} is a variation on the normalized difference index; the work of Gitelson et al. (2003) in the development of a modified simple ratio is covered in the analysis of the $CI_{red\ edge}$ Index; the Datt & Maccioni indices were modified three band difference ratios; the use of a correlogram allowed for the systematic analysis of single bands as indices, and doubled as a method of establishing a model baseline for comparison. The following sections detail how these index metrics were determined.

2.4.3.1.1 Baseline

Correlograms visualize components of the spectral and leaf biochemistry relationship, and graphically represent the results of a statistical analysis in relation to each waveband position, with the statistical performance metric of interest on the Y-axis, and the wavelength on the X-axis. Depending on the statistical metric diagrammed, correlograms capture various statistical trends along the spectrum. For example, Steele (2007) examined the RMSE to identify areas of low error consistency, while other studies used correlograms to determine spectral areas corresponding to maximum and minimum statistical variables. Typically, correlograms report either correlation or regression performance metrics. Gitelson, Keydan, and Merzlyak (2006) modeled the RMSE of linear models, while (Blackburn, 1999) modeled coefficient of determination (R^2) values for exponential models, and Quinn (2010) modeled the Pearson's correlation coefficient. Correlation was the test statistic chosen in this case, as it generated a critical value to identify thresholds for relevant performance metrics, isolating useful regions.

A correlogram identified band locations with the strongest association between spectra and pigment content. It is acknowledged the identification of significant correlation does not explain or imply causation in the dataset. The established waveband was regressed to assess and determine a causal relationship between the two variables. It functioned as a baseline against which other spectral metrics could be compared. In theory, more complex reflectance metrics should improve on the explanatory power of a single band model. Accordingly, in most instances only metrics that outperformed the single band selected through the correlogram process were considered for this study.

The correlogram was generated by taking the reflectance values, or some variation thereof (e.g. first derivative reflectance), for every destructive sample and correlated them

with the associated destructive chlorophyll pigment content of those samples. This was performed sequentially at every wavelength, and resulted in performance metrics for the entire spectrum, which were graphically modeled. Xue and Yang (2009) demonstrated there was little difference in the correlation results when considering chlorophyll in terms of units of fresh weight or area. Rather, the type of analysis used was the more influential factor. With that in mind, fresh weight was used to determine the Pearson's correlation coefficient.

A critical value was calculated as the threshold of significance, where the existence of a relationship can be established with confidence. For the correlogram, the critical value threshold was established to be ± 0.195, for n-2 degrees of freedom, $\alpha = 0.05$, and shown in Figure 3.8 as the dashed horizontal lines. Pearson's Correlation was chosen over Spearman's, as the chlorophyll content was transformed previously to obtain a normal distribution frequency, and therefore appropriate.

2.4.3.1.2 Datt & Maccioni

The Datt (1999) and Maccioni, Agati, and Mazzinghi (2001) indices (Table 2.2) were cited consistently as strong estimators of chlorophyll (le Maire et al., 2004; Sims & Gamon, 2002). Both indices used three separate wavebands from three distinct regions of the spectrum. The equations to determine the index values were

$$\text{Datt Index} = (R_{850} - R_{710})/(R_{850} - R_{680}) \quad (9)$$

and

$$\text{Maccioni Index} = (R_{780} - R_{710})/(R_{780} - R_{680}), \quad (10)$$

where *R* is the reflectance at the given waveband location denoted by the subscript number. The foundation of these formulas is a modified simple difference ratio adapted

for three bands (Datt, 1999; Girma et al., 2013; Le Maire et al., 2004). The two equations differed in the band denoting the NIR region: 850 in Datt and 780 in Maccioni.

The Maccioni equation was derived from the Datt index, and modified because of sensor range limitations. Maccioni et al. (2001) agreed with the principles of the index, but questioned the effectiveness of the sampling strategy initially presented by (Datt, 1999); therefore, the model was on single leaves instead of leaf stacks. Although often reported independently in the literature, these two indices were considered as a unit for this study, due to their mathematical and spectral similarities.

The Datt index aimed to isolate pigment subtleties that added model robustness to a range of species, and at the complex canopy-scale (Main et al., 2011; Datt, 1999). The index was based on Beer-Lambert absorption principles related to foliar biochemicals in vegetation, and included the compounding effects of internal diffuse light scattering and surface specular light scattering (Datt, 1999). This index was found to be insensitive to scattering influences, as it performed equally well as regardless if the spectra were corrected for scattering, which allowed for better characterization of biochemical absorption.

There were two stages of corrections within the index equation. The first was the mathematical determination of the difference between two sets of reflectance bands, which accounted for the additive baseline shifts that occurred from varying parameter conditions. This was determined for both the numerator and denominator. Baseline shifts manifest as changes in spectra magnitude (signal strength) (Fadock, 2011). The second correction was the ratio of the two differences, which removed the multiplicative effect of

the reflectance due to the scattering (Datt, 1999). By accounting for these two factors, the index was a function of solely pigment absorption (Le Maire et al., 2004).

The three bands used in this index were 710, 680, and 850nm. The latter provided a chlorophyll contrast; 850nm is located in the (middle) NIR region where plant reflectance is a maximum, and is insensitive to pigment absorption. For vegetation, the NIR has a higher spectral return than the visible region, and therefore provides a positive difference value between the effect of no pigment influence and strong absorption (when used as the numerator in a ratio). The comparison of a band with no pigment effect to bands of known pigment influence determined the extent of the isolated biochemical influence. This fundamental principle is employed in many indices (Datt, 1998; Murphy, Tolhurst, Chapman, & Underwood, 2005). The remaining two bands used in the index provided two discrete regions to gauge pigment absorption for the mathematical difference determination. 710nm was identified as the band position of maximum (linear) correlation to chlorophyll (Datt, 1999). 680nm is in the established band location of maximum chlorophyll absorption (Blackburn, 1999). While both these positions are related highly to chlorophyll, they differ in that the 680nm band saturations at low chlorophyll content, while the red edge shoulder does not share that limitation. The additive shift spectral component was corrected from the difference between 850nm and 710nm (scattering and RE absorption), and 850nm and 680nm (scattering and saturated maximum absorption). The ratio that resulted was a direct function of the pigment absorption, characterized by two distinct areas of the spectra (red well and RE) without a scattering factor. This index utilized more spectral information within the visible and near infrared region that pertained to pigments than its simple ratio counterparts (Datt, 1999).

2.4.3.1.3 $CI_{RED\ EDGE}$

This index was the culmination of several developmental iterations initiated by Gitelson et al. (2003), with its origins based on reciprocal reflectance spectra. The $CI_{Red\ edge}$ index was noted in Haboudane et al. (2008) as one of the top chlorophyll predicting indices. The waveband positions have changed over time (particularly red edge positioning), but this was attributed to changes in vegetation type and sensor specifications. While the $CI_{Red\ edge}$ index currently appears as a simple ratio (equation 13), it is described more accurately as a modified difference index. This is seen in an early iteration of the equation (Gitelson et al., 2003):

$$\frac{1}{(R_\lambda)} - \frac{1}{(R_{NIR})}, \tag{11}$$

where R_λ and R_{NIR} are inversed reflectances at the locations denoted in the subscript. While the log of the inverse is considered a pseudo-absorption or apparent absorbance (Blackburn, 2007; Xue & Yang, 2009), in the $CI_{Red\ edge}$ index the absorption feature is characterized more completely as the difference between two reciprocal reflectances. The first reflectance falls within the absorption feature and the second outside, and ideally both sharing similar backscattering influences (Roberts, Roth, & Perroy, 2011). The relationship was linearized by the multiplication of the NIR reciprocal reflectance, which compensated for brightness variations that were backscatter-dependent (Gitelson et al., 1996). This equation is comprised of three components: a band maximally sensitive to chlorophyll, as well as any other biochemicals influencing the absorption feature; a band sensitive to other biochemicals but minimally to chlorophyll; and a band which characterizes the effects of scattering due to leaf structure. This yielded Gitelson et al. (1996)'s equation

$$[(R_\lambda)^{-1} - (R_{NIR})^{-1}] * R_{NIR} = \left(\frac{R_{NIR}}{R_\lambda}\right) - 1, \quad (12)$$

where R_λ is the reflectance of the red edge, and R_{NIR} is a reflectance value that falls within the near infrared region. Mathematically, when the brackets are expanded, R_{NIR} is divided by itself, resulting in the *-1* component in equation 12. The subtle difference of origin distinguishes itself from other simple ratios.

Steele (2007) developed this index specifically for grape leaves, making it a suitable comparative study. The final index (equation 13) compares signals of a region highly sensitive to chlorophyll pigments with a region that is specifically not influenced by chlorophyll absorption characteristics (Gitelson & Merzlyak, 1996).

$$CI_{\text{red edge}} = (R_{NIR}/R_{\text{red edge}}) - 1 \quad (13)$$

A critical component of this index was determining the individual equation parameters, R_{NIR} and $R_{\text{red edge}}$. Originally intended for broadband, the authors identified homogenous regions from which one specific waveband, regardless of the spectral resolution, could represent the equation parameters. The selection of these two positions is expanded upon in the results section.

2.4.3.1.4 NDVI

Arguably the most prolific index, the normalized difference vegetation index (NDVI) is considered a vegetation structural index, in that it principally is related to biomass and canopy structure (Stagakis, González-Dugo, Cid, Guillén-Climent, & Zarco-Tejada, 2012). Originally developed using broadband satellite data by Rouse, Hass, Deering, and Schell (1973), the NDVI index identifies vegetation within a multispectral image. The image classification index is given by the equation as

$$NDVI = [R_{NIR} - R_R]/[R_{NIR} + R_R], \quad (14)$$

where R_{NIR} and R_R are reflectance values at the two subscripted spectral regions (Rouse et al.,1973). Index values range from -1 to +1, and are used to separate as many materials as possible from vegetation in a spectral image. For example, water returns negative values, as the R_R returns a stronger signal than the R_{NIR}. Bare soils have a fairly neutral value (0 to 0.2), but vegetation has a higher positive return, especially as the density of canopy increases. Many current iterations of the index focus on emphasizing minute variations in higher vegetation range values.

Mathematically, the index is a simple ratio of R_{NIR} over R_R that utilizes the prominent red edge characteristic of vegetation. The difference calculation in the numerator is responsible for the positive or negative value of the index, and works to identify and isolate the vegetation response. The summation calculation in the denominator counters the effects of spectral inconsistencies due to scaling factors. The index ratio calculation normalizes in two instances: First, pigment absorption of the red well is normalized to the cellular structure and diffuse internal scattering of the NIR. Second, the reflectance is normalized to sensor, season, and site location differences through the relative ratio of the difference (numerator) to the sum (denominator), at two discrete regions of the spectrum. Although this approach enhances pigment absorption and minimizes background effects, the NDVI is limited fundamentally due to the rapid saturation at low chlorophyll contents, a characteristic of the red well (Serrano, Peñuelas, & Ustin, 2002). NDVI is known to saturate at high vegetation densities or multi-layered canopies (Mutanga & Skidmore, 2004; Thenkabail, Smith, & De Pauw, 2000; Wu et al., 2008).

To counter the saturation limitation, alterations and additions to the original index equation have been proposed. As previously established, the modified normalized difference index (mND_{705}) was selected for this study. This iteration emphasized the NDVI as a foliar pigment estimator; to avoid saturation, the 705nm subscript identified an off-peak band pertaining to foliar chemistry, rather than the previous red well. The equation by Sims and Gamon (2002) is given as

$$(R_{750} - R_{705})/(R_{750} + R_{705} - 2R_{445}), \quad (15)$$

where *R* is the reflectance at the wavelength denoted in the subscript. It has been demonstrated that off-peak bands do not saturate as quickly (Le Maire et al., 2004). A band with a weighting factor was added to further correct scattering; it accounted for leaf surface specular scattering, which can increase reflectance magnitude across the visible portion of the spectrum (Sims & Gamon, 2002). The strong absorption from biochemical pigments (both chlorophylls and carotenoids) in the blue region results in very low reflectance levels, meaning that any significant reflectance magnitudes are attributed to surface scattering.

2.4.3.2 Derivatives

This section outlines eight derivative metrics regressed with chlorophyll for its remote non-destructive estimation. The wavelength position of the maximum slope change, the Red Edge Inflection Point (REIP), is the traditional derivative metric well established within the literature (Baranoski & Rokne, 2005; Mutanga & Skidmore, 2007; Sims & Gamon, 2002). Reported less often is the derivative reflectance value at this inflection point (defined here as MaxD). Cho and Skidmore (2006) presented a new method of determining the REIP from the first derivative: the convergence point of two

straight lines that flank the red edge (dashed lines in Figure 2.8). The REIP metric value is a ratio of the slopes and intercepts of the two straight lines. The interpolated lines connected an anchor point along the X-axis (680 and 760nm for each line, respectively) to the tangential position along the derivative slope. The wavelength location of the tangential-line intersections provided the reflectance variable of interest. Originally developed to investigate nitrogen, the linear extrapolation approach was a novel use of derivative data, and its feasibility as a chlorophyll estimator was assessed (Cho & Skidmore, 2006). It was hypothesized that this method could be applied to chlorophyll, as a strong relationship exists between chlorophyll and nitrogen. The metric is referred to as AltREIP in this study.

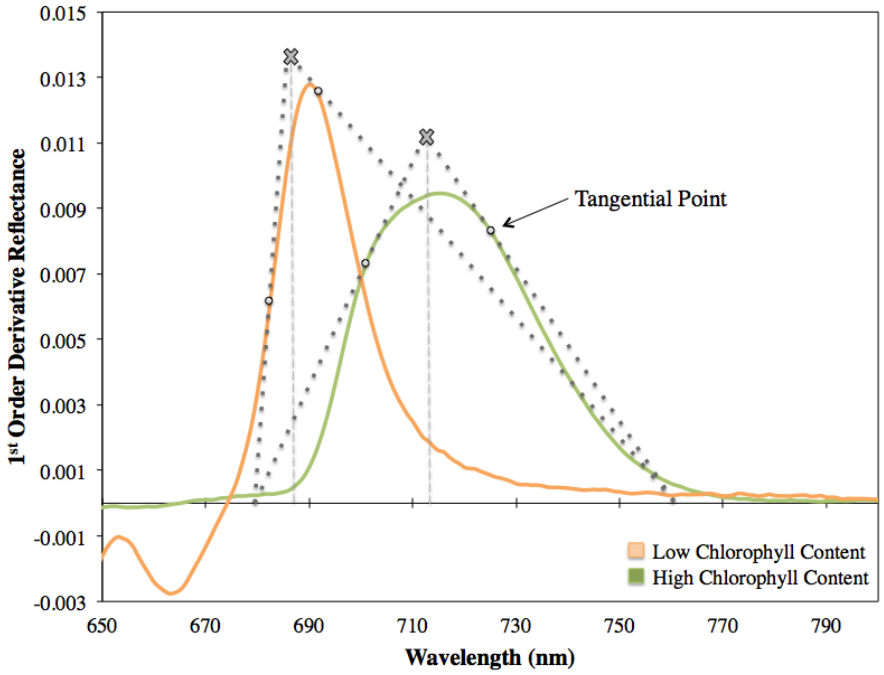

Figure 2.8: Example of two varying first order derivative spectral responses, based on maximum and minimum chlorophyll content. Grey lines are the extrapolated lines of the derivative derived connecting one point along the derivative curve to one an anchor point on the X-axis. Their intersection point (grey X) is the location of the AltREIP. Adapted from (Cho & Skidmore, 2006)

A correlogram reporting the Pearson's correlation coefficient between the 1st order derivative reflectance and destructive pigments determined the single waveband with the strongest degree of association used as a Single Band index (SBI). The method of obtaining the correlogram was outlined previously (section 2.4.3.1.1), the only difference was the first order derivative reflectance was used instead of the unaltered reflectance.

Le Maire et al. (2004) completed a pivotal comprehensive investigation of derivative spectra and chlorophyll estimation. While the study had a simulated component, emphasis was placed on the experimental dataset results. It covered more traditional methods of derivative work (e.g. REIP), as well as three novel derivative indices: The Double Difference Index (DDI), the derivative modified normalized ratios of differences (R'mND) and the derivative modified simple ratios (R'mSR) index.

The equation of the DD index was given as

$$DD = (R_{b+\Delta} - R_b) - (R_{a+\Delta} - R_a), \quad (16)$$

where *R* was the unaltered reflectance values (Le Maire et al., 2004). Derivative data was used to identify four positions that capture the difference in the derivative slope on either side of the red edge; *a* and *b* were two anchor points on the shoulders of the derivative peak (dashed lines in Figure 2.9), and the distance to the two ancillary bands locations was the same (Δ). Once the distance was determined for one side, it was applied to the other side (giving b+ Δ and a+Δ). Three different methods of determining *a* and *b* were discussed in the results section.

Developed with the reflectance data, the application of the following two indices to derivative data was a new approach. The equation for R'mND Index was

$$\frac{(R'_a - R'_b)}{(R'_a + R'_b - 2R'_c)}, \quad (17)$$

where *R'* denotes first derivative reflectance values, and *a, b,* and *c* represent waveband locations. This index equation, a modified normalized difference class of index, is found in Sims and Gamon (2002), and denotes *a* as the reference spectra normalized to *b*, the spectra associated with the foliar characteristic of interest. The weighting factor constant, 2*c*, modified the index to account for specular surface reflectance differences between leaf types. The waveband location identified as insensitive to surface scattering was 445nm, and 750nm and 705nm were determined for *a* and *b* respectively (Sims & Gamon, 2002). Le Maire et al. (2004) proposed the use of 722, 699, and 502nm for *a*, *b*, and *c*, to adjust for the used of derivative values instead of reflectance. The shorter derivative wavelengths had significant amounts of noise (negating the use of 445nm), and the longer wavelengths only captured the far tail of the red edge peak.

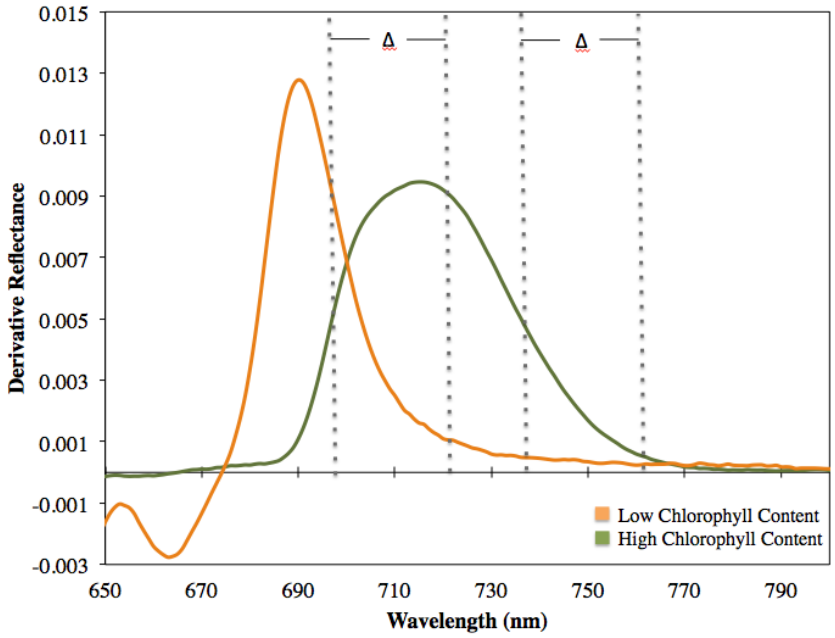

Figure 2.9: An example of two derivative spectra in the red edge region, and the visualization of the variables used to determine the DD Index. Adapted from (Le Maire et al., 2004)

Similarly, the equation for the R'mSR Index was

$$\frac{(R\prime_a - R\prime_c)}{(R\prime_b - R\prime_c)}, \quad (18)$$

where *R'* was the first derivative reflectance values, and *a*, *b*, and *c* denote waveband locations. These locations captured the same features as the previous equation. In this equation iteration, the constant *c* was subtracted directly from the reference spectra, *a,* and the spectra of foliar biochemical interest, *b*. The same derivative band locations as the mND index were used, as they determined them as the most suitable waveband candidates through a calibration and validation method (Le Maire et al., 2004). This provided the opportunity to assess of the robustness of these indices for grape leaf data.

2.4.3.3 Continuum Removal

While indices presented previously relied on individual waveband reflectance magnitude differences to determine spectra differences, the continuum removal method utilized the entire reflectance absorption shape (Kokaly et al., 2003). This took advantage of an underutilized characteristic in the fundamental composition of reflectance data. Ordinary least squares regression was used to assess the continuum metrics most closely related to chlorophyll content, and identified the most appropriate continuum metric for viticulture research.

The analysis began with the identification of the region of interest that characterized the absorption feature, in this case chlorophyll *a* (Visintini, 2010). The isolation of the absorption feature is required to study it (Huang et al., 2004; van der Meer, 2004). 550-780nm encompassed the region of the electromagnetic spectrum from the green peak to the red edge plateauing in the NIR. There is no established methodology for determining the exact end points of the continuum line. Although an empirical approach was used for

this study, the final determinates were subject to alterations to accommodate data spread. Without such, there were problems with the practical application, as further analysis was impeded if the continuum line crossed the reflectance line and yielded negative values.

Kokaly et al. (2001) indicate that endpoints capturing the shoulder of a feature were preferable to extreme extents. For the first endpoint, or the left side of the boundary feature, the green peak region of the electromagnetic spectrum was of interest. The green peak maxima was located with reflectance spectra, and from the derivative data, the point of intersect where the derivative crossed the X-axis (which indicated a change from positive to negative slope). They were determined to be 563nm and 556nm, respectively. The average of the two provided the final wavelength position for the first continuum endpoint. The second endpoint, or the right side of the boundary, encompassed the red edge region. The derivative spectrum was utilized to determine the REIP for all samples. To this end, the average REIP location for all samples was determined to be 720nm, as a fixed point captured the comparative changes between samples. These endpoints complement the work of Kokaly et al. (2007) who determined the ranges 512-542nm and 737-767nm for both forest and non-forested environments. Alternative methods for endpoint selection include more current work of Kokaly (2011), who developed the software module PRISM for the geospatial image analysis software ENVI, which determines non-static endpoint unique to each sample. Similarly, Sanches and Filho (2014) found that adopting a moving end-point detection approach improved characterization of induced stressed grass plant species. However, these methods are limited to either imagery data (with ENVI plug-in) or custom software.

With boundary points of 560nm and 720nm, the slope (s') and y-intercept (y') value of the continuum line was calculated for each destructive sample with the equations

$$s' = (R_{720} - R_{560})/(\lambda_{720} - \lambda_{560}) \quad (19)$$

and

$$y' = R_{560} - (s' * \lambda_{560}). \quad (20)$$

R was the reflectance value and λ was the wavelength position at the given band. The linear function of the continuum enabled the generation of simulated point values along the continuum for every corresponding wavelength measurement taken with the ASD spectrometer.

This approach resulted in two reflectance values for every wavelength: a laboratory measured reflectance (green and orange in Figure 2.10), and a second simulated reflectance of the spectra without the influence of the absorption feature (black line in Figure 2.10a). The division of the measured reflectance by the simulated continuum generated a continuum removed (CR) value for all 160 wavelengths.The band depth (BD**)** is the continuum treated as reflectance rather than absorbance, and is simply the CR subtracted from 1 (Quinn, 2010). The maximum band depth (MBD) value for each sample was identified after the transformation as the largest value in the spectral range. Each continuum-removed band was normalized to the MDB to obtain the normalized band depth (NBD) metrics. Making all values between 0 and 1 allowed for comparison to other datasets (Kokaly & Clark, 1999). In the case of airborne datasets, normalization also minimizes external confounding effects, such as atmospheric scattering and topographic influences (Kokaly & Clark, 1999; Mutanga et al., 2005).

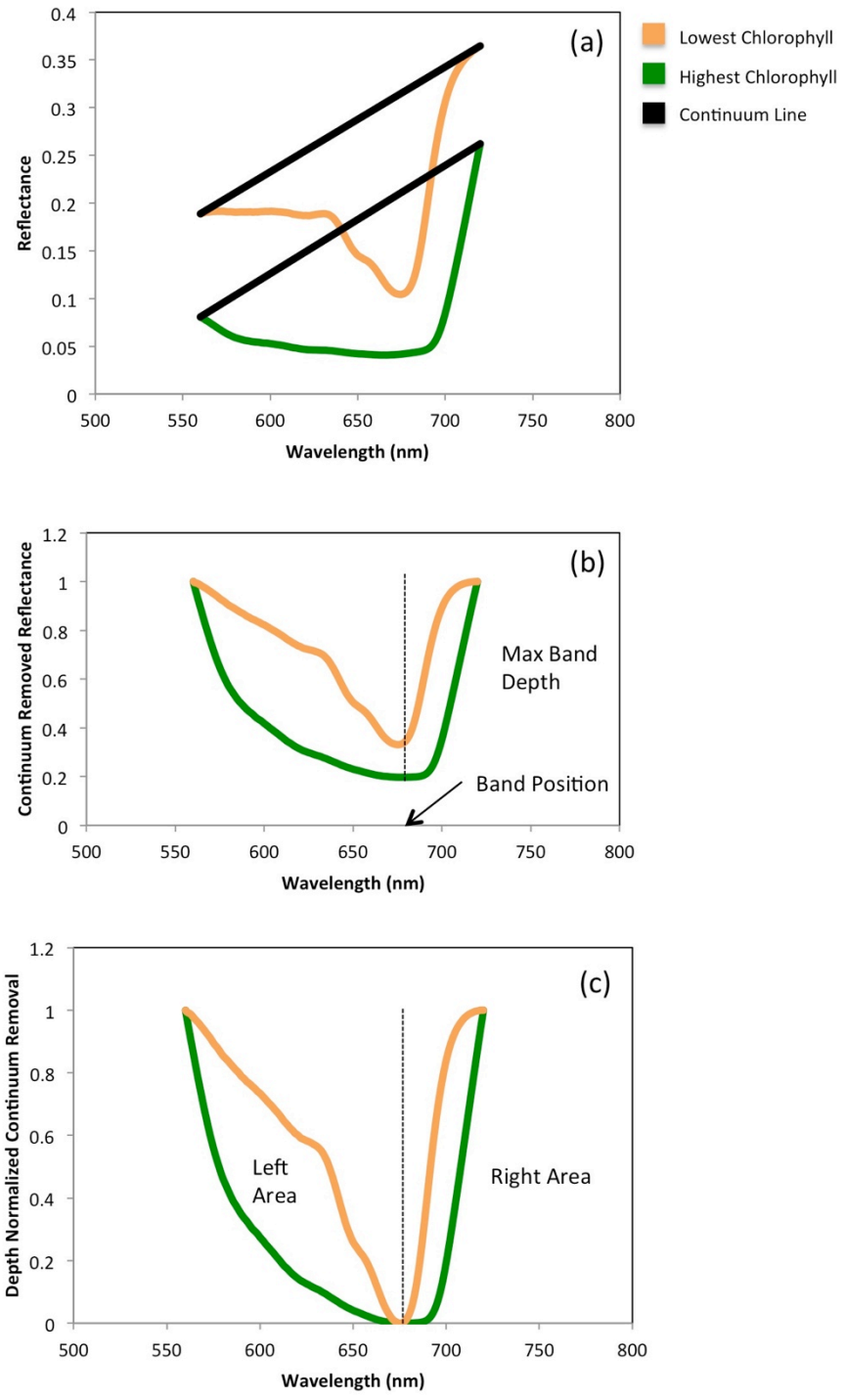

Figure 2.10: (a) the isolated reflectance spectra and corresponding continuum for the chlorophyll absorption feature; (b) continuum removed reflectance spectra, which allows the identification of the depth of the maximum absorption of all samples; (c) samples are normalized to allow for comparison and analysis. Areas under the continuum area determined**.** Adapted from (Girma et al., 2013).

In addition to these metrics, a further set of variables calculated the area created under the normalized continuum curve from the pigment absorption. Total Area (TA) was the summation of individual areas between each waveband, and were in one-nanometre increments. Area beneath the NBD was calculated using the trapezoid method (Quinn et al., 2012), which added the depth of two sequential wavelengths and divided by two.

Individual Left (LA) and Right (RA) areas were established using the location of the MBD as the separation point (Figure 2.10c). The mathematical ratio between the left and the right area captured the feature symmetry of the continuum removal (AS). The final metrics generated for analysis are found in Table 2.6. A parameter not included in this study was the width of the absorption feature, characterized as the full wavelength width at half the maximum depth. The change detected by this metric was represented in the area measurements, and thought superfluous (Pu, Ge, Kelly, & Gong, 2003; Rodríguez-pérez, Riaño, Carlisle, Ustin, & Smart, 2007).

Table 2.6: List of the Continuum Metrics generated and used in the analysis of this thesis.

Metric	Abbreviation
Maximum Band Depth	MBD
Maximum Band Depth Position (nm) where the MBD occurred	MBDP
Continuum Removed Total Area	CRTA
Right Area/Left Area	CRRA/CRLA
Left and Right Area Symmetry	AS

Chapter 3 Results

The results below are a series of univariate ordinary least square regressions with the common dependent variable, destructive chlorophyll measurements. Reflectance metrics from indices, derivatives, and continuum removal were generated (detailed in Section 2.4.3), and were regressed to report the varying results against each other. The same dependent variable was used for all regressions so the model error assessments were comparable. The best metric was selected and rigorously assessed through data splitting and jackknifing validations to compare the predicted values generated from the calibration model with an independent data subset. This study incorporated both data-splitting and a data resampling approach for a more thorough validation analysis, similar to Mutanga et al. (2004) for a savanna assessment study who combined bootstrapping with cross-validation techniques.

3.1 SPAD

3.1.1 Calibration

The chlorophyll meter dataset, having only one metric, could be immediately split into calibration and validation. For the training dataset, a power function captured the relationship between SPAD values and chlorophyll expressed as a fresh weight concentration, coefficient of determination (R^2) = 0.89, standard error of the estimate (RMSE) = ±3.40 μg/g. The transformed chlorophyll fresh weight was previously established as normally distributed, and therefore suited to regression modeling. The validation final equation was

$$\sqrt{\text{Chlorophyll Content (}\mu\text{g/g)}} = 2.85*\text{SPAD}^{0.694}. \quad (21)$$

The residuals from this model had a mean of zero and a skewness of -0.361, attributed to one sample, which occurred three standard deviations below the zero axis. This did not affect the overall normality of the distribution, as the residuals passed the K-S and S-W tests ($p = 0.20$ and 0.59, respectively). When a second iteration of the regression was performed, a quadratic function best characterized this relationship, and while the additional *x* variable only increased the coefficient of determination by 1%, it was a significant F-stat change ($p < 0.05$), $R^2 = 0.90$, RMSE = ±3.50 μg/g. In addition, the second order polynomial produced a more uniform spread of the residuals when graphed against the dependent variable, and were distributed normally with a mean of zero. The final equation was

$$\sqrt{\text{Chlorophyll Content}}\ (\mu g/cm^2) = 0.227 + 0.151*\text{SPAD} + (-0.001*\text{SPAD}^2) \quad (22)$$

There were three identifiable results when compared with the literature. The first was the shape of the regression models; both calibration models (area and fresh weight) exhibited a general logarithmic growth trend (Figure 3.1), as the dependent chlorophyll estimate (Y-axis) values plateaued and became invariant with increased independent SPAD measurements (X-axis). Conversely, in studies with a variety of species, the relationships tended toward exponential growth trends, indicating the SPAD (X-axis) values were the variable to become invariant as the chlorophyll (Y-axis) increased (Bannari, Khurshid, Staenz, & Schwarz, 2007; Cerovic et al., 2012; Markwell et al., 1995; Uddling et al., 2007).

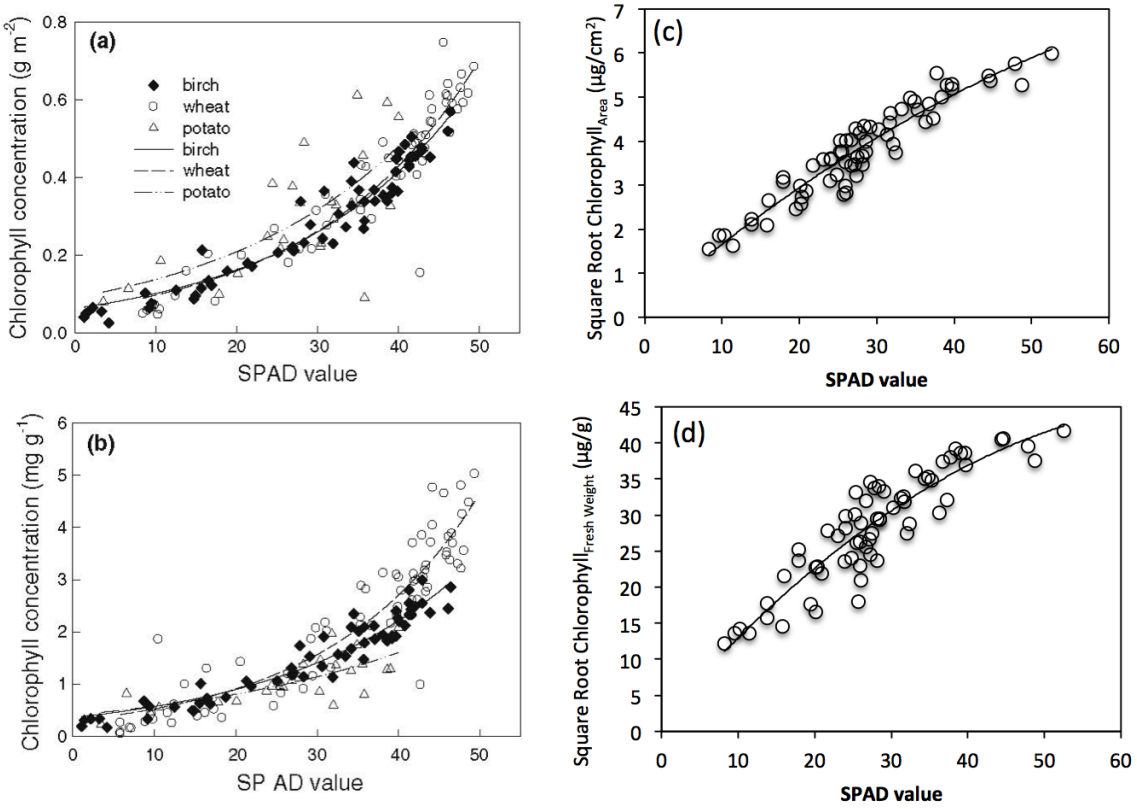

Figure 3.1: Comparison of scatterplots between the results of Uddling et al. 2007 (a & b) and that of this study (c & d). The top figures compared chlorophyll expressed in terms of area, and the bottom were chlorophyll in units of fresh weight. Although the units for the studies differed, they are similar ranges for both pigment and SPAD values. The transformation of the dependent variable contributed to the differing model shapes.

This logarithmic scatterplot shape was the result of the transformed value of the chlorophyll measurements (Figure 3.1). The transformation overcorrected as a method for the non-normality of the dependent variable. A modified custom exponent would have adjusted to the correct degree, but would hinder the robustness of the predictive modeling and render it more dataset-specific. Results obtained by Lin et al. (2015), who also performed a transformation on the dependent variable, had a similar distribution. However, Broge and Mortensen (2002) performed a logarithmic transformation of chlorophyll data in favour of an exponential relationship between SPAD and leaf chlorophyll. Linear relationships were also identified (Campbell et al., 1990; Ruiz-

Espinoza et al., 2010). The work of Azia and Stewart (2001) identified chlorophyll by area to have a quadratic model, as was the case in this study. Unlike Azia and Stewart however, this study determined chlorophyll expressed as fresh weight to have a power model, whereas they optimized with a quadratic model. The diversity of results in the literature indicate that, while the results of this study did not emulate the frequently reported exponential relationship, relationship trends depend largely on phenological stage, species variety, extraction methods, data range, sampling size, data processing, and other factors.

The second result of interest was model performance. There was agreement in the literature regarding transmittance measurements; chlorophyll expressed per unit area yielded stronger results than by fresh weight (Azia & Stewart, 2001; Ruiz-Espinoza et al., 2010; Uddling et al., 2007). That said, results varied with such factors as independent or grouped species, and whether chlorophylls were investigated individually or as total pigment content. The present results concurred, as the SPAD was able to account for 8% more variation in chlorophyll measurements in terms of area than the fresh weight counterpart. Overall the model performance was well situated in the literature, ranging from $R^2 = 0.97$ (Netto et al., 2005) down to 0.46 (Ruiz-Espinoza et al., 2010; Uddling et al., 2007).

The third general finding concerned the identification of a chlorophyll sensor sensitivity limit of 300mg/m^2 for grapevines (Steele et al., 2008). The authors determined the SPAD instrument could accurately predict chlorophyll content below this threshold, however, not well in grape leaves exceeding this chlorophyll content. This threshold was converted to the appropriate units and visualized in Figure 3.2. In this senescent dataset,

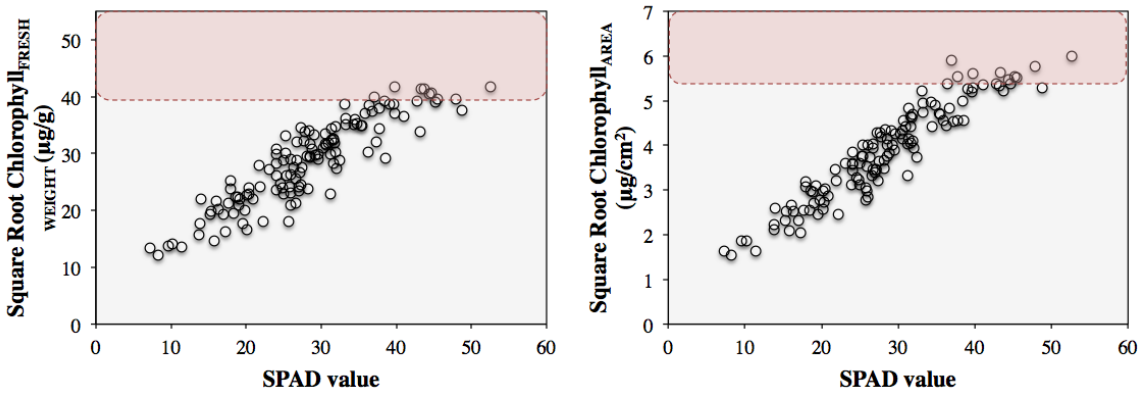

Figure 3.2: The red area indicates the approximate limit of the SPAD's ability to accurately estimate chlorophyll, as determined by Steele (2008a). When applied to this dataset, it indicates little of the dataset would be affected by this limitation.

very few samples surpassed this threshold, so in theory should not be greatly affected by this limitation. This threshold is addressed at more length in the discussion.

3.1.2 Validation

The calibration equations were used on the validation dataset for an internal model assessment, and the results were squared to yield chlorophyll content in both fresh weight and area units. These values were compared with the extracted pigment content of the corresponding samples. The results yielded strong linear relationships for both units of measurements.

Figure 3.3a showed a larger deviation from 1:1 ratio line (estimate:measured) for the area measurements, which resulted in a underestimation of chlorophyll at higher content levels. This resulted in the upper range of estimated chlorophyll value terminating at 25.3 µg/cm^2, while the actual maximum was 34.8 µg/cm^2. Despite this, both the fresh weight and area models yielded high coefficients of determination, R^2 = 0.80, RMSE = ±158.43 µg/g, and R^2 = 0.86, RMSE = ±2.15 µg/cm^2, respectively.

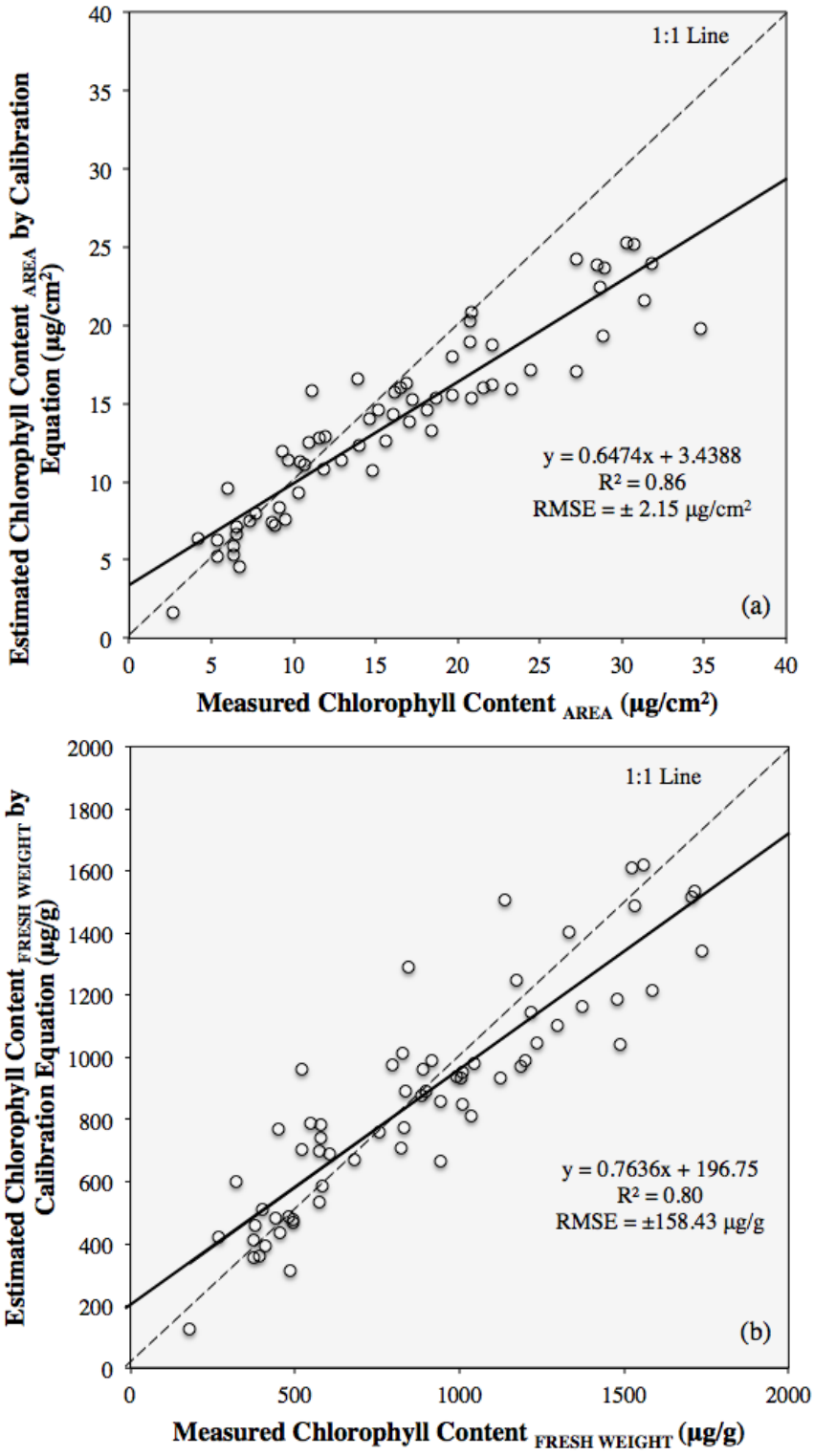

Figure 3.3: A power and quadratic equation were used in the generation of calibration models of chlorophyll (a) area and (b) fresh weight, respectively. SPAD values were then used to generate estimates for the validation dataset, and the estimates were compared to the destructively determined measurements. Chlorophyll by area deviated more than the fresh weight model.

Comparatively, Lin et al. (2015) found large deviations between their calibration and validation datasets when controlling for water stress in camphor tree leaves; models built on hydrated leaves were not applicable to wilted leaves. In terms of factors controlled within this dataset, the stratification of chlorophyll content ensured that the pigment content was not lower in the calibration modeling compared to the validation, and therefore not a major source of deviation.

3.1.3 Jackknifing

This resampling approach was performed to further modeling analysis, while reducing bias from the dataset (Lawrence & Labus, 2003). The overall accuracy of the SPAD as a non-destructive chlorophyll estimator was higher for chlorophyll expressed per unit of fresh leaf area than per unit of fresh weight, $R^2 = 0.87$, RMSE = ±183.43 μg/g, ranging from 148.94 to 1736.89 μg/g, and $R^2 = 0.79$, RMSE = ±2.98 μg/cm^2, ranging from 2.37 to 35.82 μg/cm^2. These jackknifed regression results are less sample dependent than the previous calibration model, and better reflect the grapevine population of interest (Meeuwig & Peters, 1996). The regression equations for the jackknife validation were

$$Chl_{predicted}\ (\mu g/g) = 0.834 * Chl_{measured} + 134.6 \quad (23)$$

and

$$Chl_{predicted}\ (\mu g/cm^2) = 0.903 * Chl_{measured} + 1.42 \quad (24)$$

for fresh weight and area measurements, respectively, as seen in Figures 3.4a and b.

The jackknife performance metrics indicated the chlorophyll meter is compatible with destructive measurements in area terms, as it accounted for more variation in pigment content than in fresh weight form. This aligned with the previous data-splitting validation results. The coefficients of determination for the area measurements were the

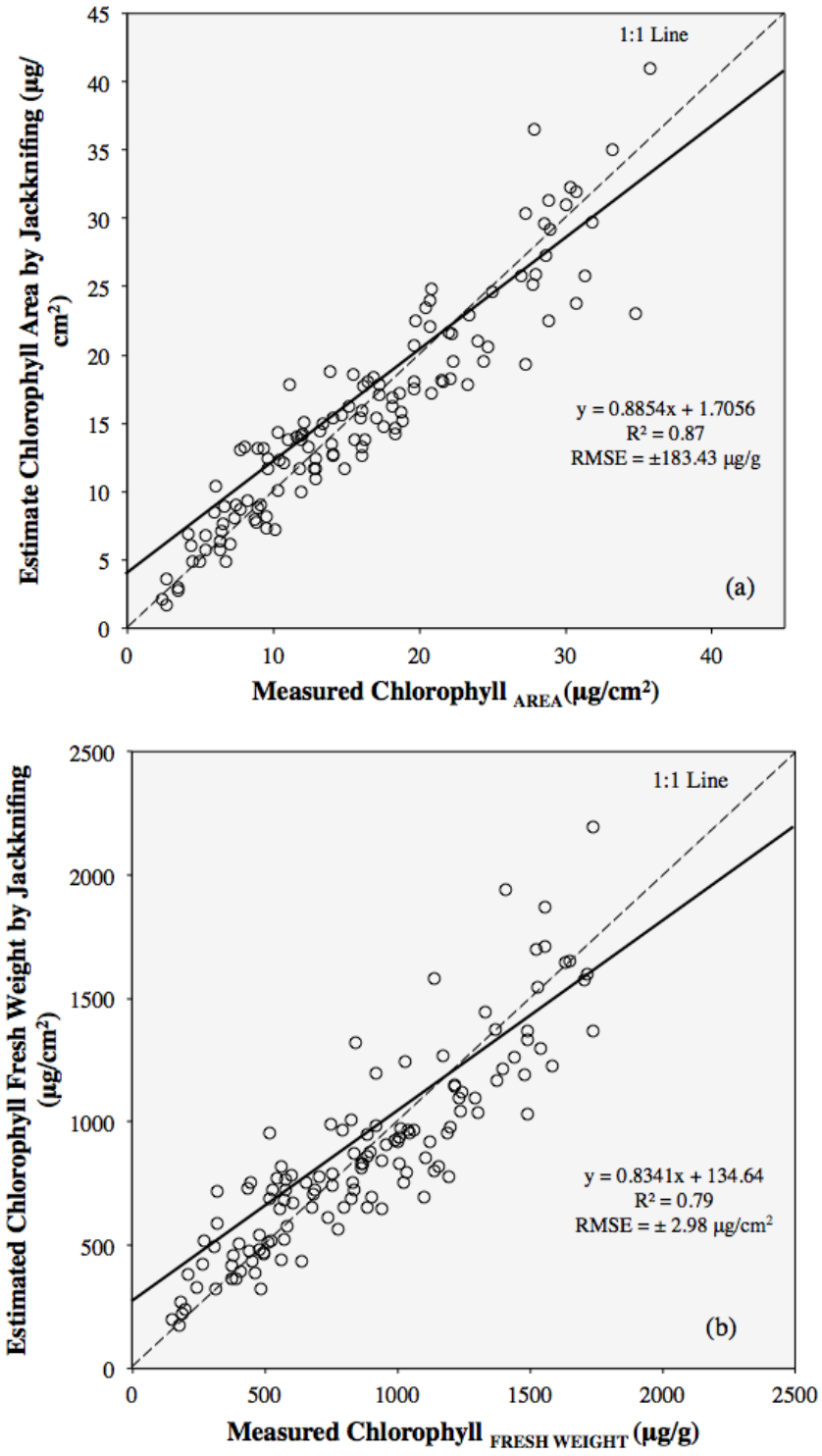

Figure 3.4: Comparison of (a) area chlorophyll estimate to actual destructive chlorophyll content for SPAD modeling jackknifing validation and (b) fresh weight chlorophyll estimate to actual destructive chlorophyll content for SPAD modeling jackknifing validation.

same between the two methods, while the fresh weight yielded slightly lower results in the jackknifing. The RMSE was larger for both jackknifed models compared to the data-splitting method. The larger error was an anticipated result as the jackknife calculations are more robust, as they are less susceptible to individual samples or bias (Novoa, Chust, Valencia, Froidefond, & Morichon, 2011). That said, the results for the SPAD estimates of chlorophyll per unit area were heteroscedastic, and consequently the estimates were less accurate at the higher chlorophyll contents (Figure 3.4a). This lends weight to the findings of Steele et al. (2008) regarding a chlorophyll meter threshold limit, but starting at a lower range. However, Le Maire, et al. (2004) showed a similar increased scedasticity in jackknifed results based on area measurements without a chlorophyll meter, suggesting it may not be a sensor specific characteristic. There was an overestimation of samples consistently throughout the fresh weight chlorophyll value range (Figure 3.4b). Samples #9, 15 and 8 were identified as isolated samples in area measurements, while samples #19, 15, 47, 58 and 63 were identified for fresh weight. These samples were not influential, but rather constitute the expected 5% of the 134 samples that fall outside the second standard deviation point of the data spread.

3.1.4 Literature Models

Two studies were chosen for comparison, as they developed multi-species models built for robust, species independent chlorophyll estimation. Uddling et al. (2007) presented fresh weight models that achieved higher coefficients of determination than this study. They determined an exponential equation

$$y = 0.0691\ e0.0459x, \tag{25}$$

where x is the SPAD value and y is chlorophyll content in g/m^2. The chlorophyll was determined and adjusted to the units reported in this study. The resulting estimates had a mean of 26.18 µg/cm^2, and a minimum and maximum of 8.99 µg/cm^2 and 74.46 µg/cm^2, compared to the known range of 2.37 µg/cm^2 to 35.82 µg/cm^2. The model performed well, R^2 = 0.82, RMSE = 4.99 µg/cm^2; however, the residuals were highly heteroscedastic (Figure 3.5). This variance inconsistency, combined with significant overestimation in pigment content, determined the published equation as unsuitable for this grapevine study.

Cerovic et al. (2012) presented a new index that amalgamated seven models from other prominent studies (Cartelat et al., 2005; Coste et al., 2010; Ling, Huang, & Jarvis,

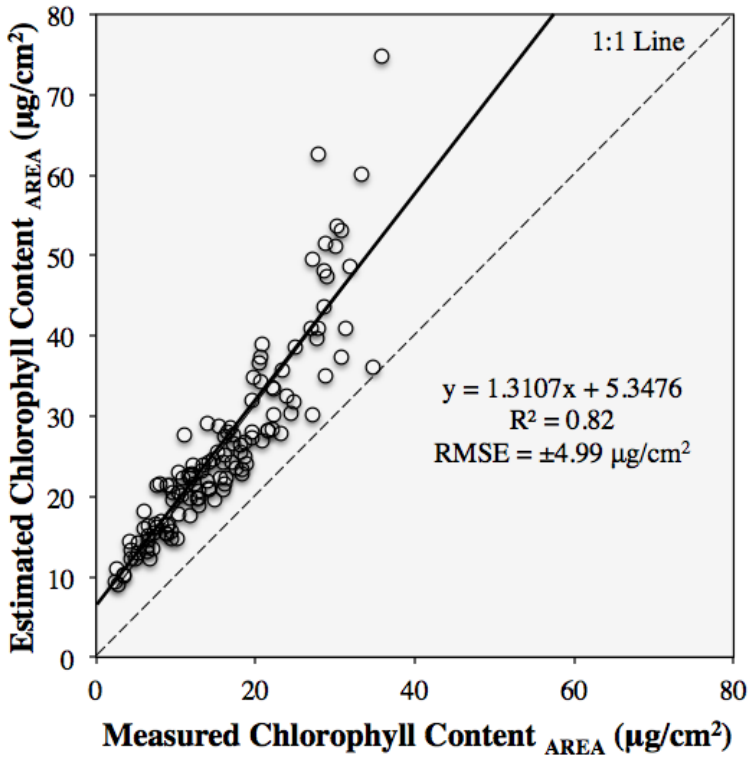

Figure 3.5: Comparing the estimated chlorophyll determined using the Uddling et al. (2007) equation with the measured chlorophyll content of this study.

2011; Marenco et al., 2009; Markwell et al., 1995; Richardson et al., 2002; Uddling et al., 2007) in addition to their own independent calibration (Figure 3.6). The equation generated chlorophyll content estimation in μg/cm^2 as

$$y = (99*SPAD)/(144-SPAD). \quad (26)$$

The regression between estimated and measured chlorophyll yielded a strong linear model, with better dispersion in the residuals than the previous Uddling et al. (2007) model, R^2 = 0.87, RMSE = ± 3.59 μg/cm^2. However, it also overestimated the chlorophyll content on a per unit area basis (Figure 3.7). The estimated range was 5.24 to 56.97 μg/cm^2, while the measured chlorophyll area values had a maximum and minimum of 35.2 and 2.3 μg/cm^2 with a mean of 15.89 μg/cm^2.

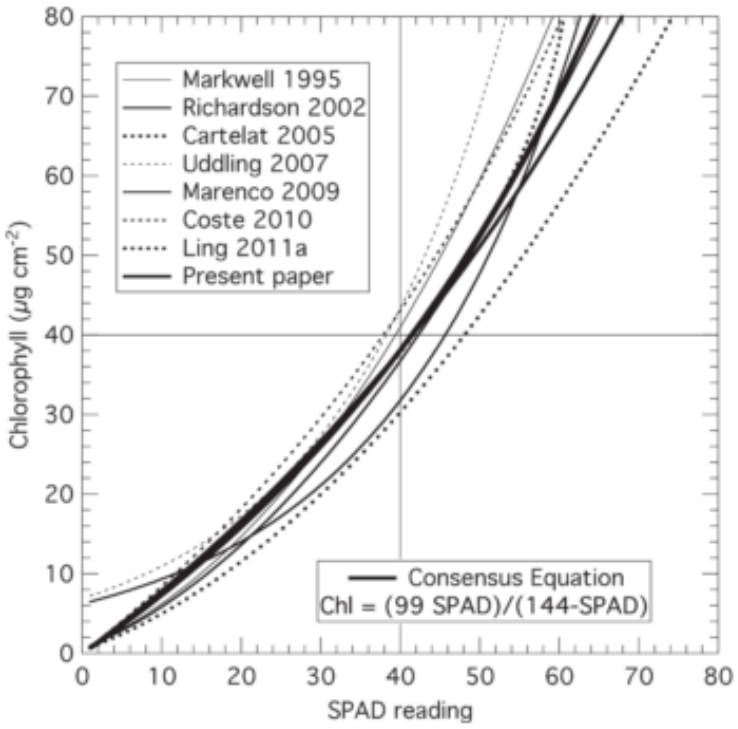

Figure 3.6: Cerovic et al. (2012)'s consensus equation that combined the details of other studies, in addition to their own calibration.

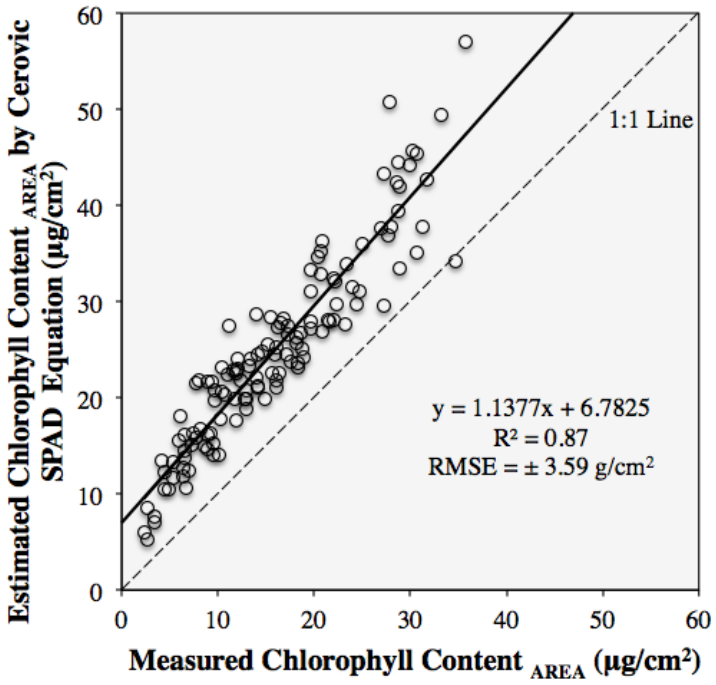

Figure 3.7: Comparing the estimated chlorophyll determined using the Cerovic et al. (2012) equation with the measured chlorophyll content of this study.

3.2 Reflectance

3.2.1 Baseline & Indices

3.2.1.1 Theoretical Expected Results

The four index metrics reported here emphasized the red edge and NIR regions of the spectrum. The spectral variations in the red edge are characterized by changes in slope drift and magnitude, which are directly influenced by strength of the red absorption feature, water content, and scattering. The indices utilized specific bands to capture these changes. The mathematical manipulations vary, but the use of a normalizing ratio was reoccurring; it normalized a spectral region affected by a particular characteristic to a region not influenced by that same characteristic. The quotient isolated the specific pigment absorption spectral response without effects that were present in both regions.

3.2.1.2 Summary

No one index stood apart in quality, as performance metrics all reported similarly (Table 3.1). There were consistent complications with residual patterns in the datasets. NDVI and Datt/Maccioni indices identified the same 3 samples as extremes (#167, 168, and 102), and the single band index located four different samples that behaved uncharacteristic of the rest (sample #141, 145, 148 and 158). However, none of these sample's features, in either chlorophyll determination or reflectance spectra, were extreme enough to warrant classification as outliers, unlike those identified in section 2.3.1. Rather, it was indicative of senescent behaviours and ranges appropriate to the dataset.

The overall results of index performance showed several features. The CI red edge, Datt & Maccioni, and NDVI indices performed equally well for chlorophyll estimation that used a variety of model functions. The highest coefficient of determination was 0.92,with a corresponding standard error of estimate of ±2.10 μg/g. The mND_{705} Index was the only index that did not require a dataset-specific band location selection; the other indices utilized performance metric correlograms or

Table 3.1: Summary of the performance metrics results for the reflectance-based index regression models.

Index Name	Coefficient of Determination R^2	Standard Error of the Estimate RMSE	Model Expressions
Single band	0.89	±2.50 μg/g	Linear
Datt & Maccioni	0.92	±2.10 μg/g	Linear
CI red Edge	0.92	±2.11 μg/g	Exponential
mND_{705}	0.92	±2.11 μg/g	Quadratic

validation methods to determine if the model could be improved and customized for grape leaf specific results. As hypothesized, all indices achieved better results than the baseline, as they utilized additional spectral information to improve on pigment estimation.

3.2.1.3 Correlogram

A correlogram reporting the Pearson's correlation coefficient between the unaltered reflectance and destructive pigments is presented in Figure 3.8. The dashed lines represent the critical value threshold, as values outside (greater than in the positive, less than in the negative) were considered significantly correlated. Band 771nm had the

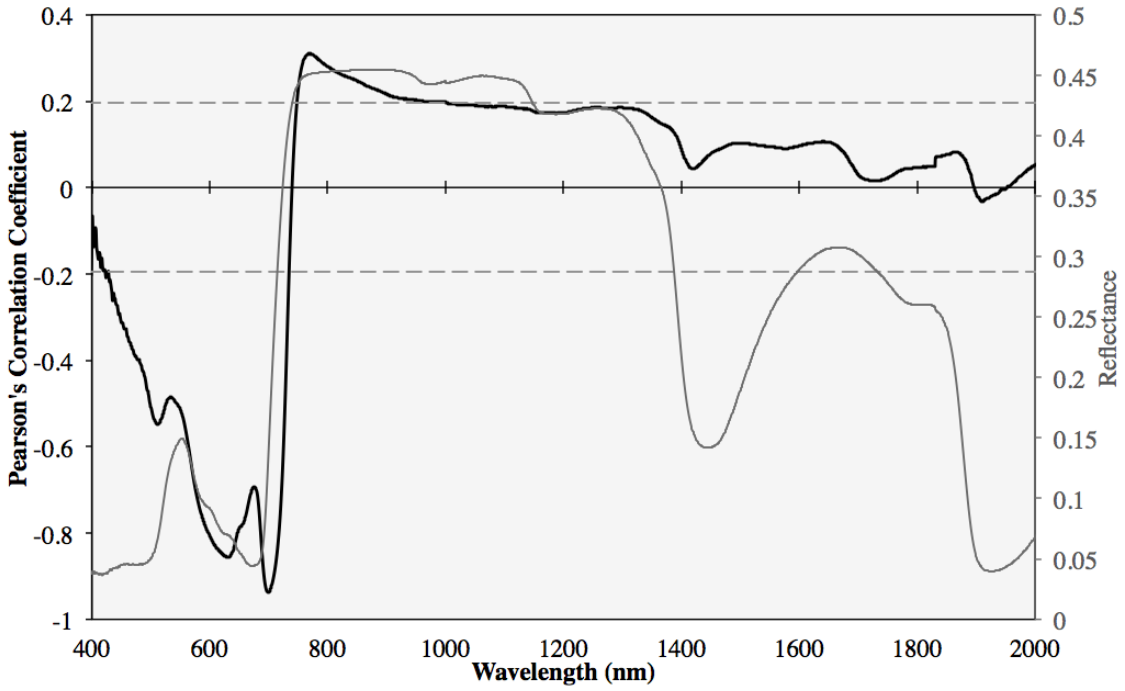

Figure 3.8: Solid black line is the correlogram: the non-destructive reflectance measurements for all samples at each wavelength correlated with destructive pigment content to yield a continuous Pearson's correlation coefficient graph. A typical reflectance measurement is included for reference (grey line). Three local negative maxima can be seen in the 400-800nm range (550, 630, and 702nm).

strongest positive linear association with chlorophyll (Pearson's correlation coefficient of 0.248), and band 702nm had the strongest negative linear association (correlation coefficient with -0.906). These locations complimented Wu et al. (2008), who identified 705nm optimal for achieving linearity when related to chlorophyll concentration. Two continuous regions were identified as related significantly to chlorophyll. The first was from 432nm to 734nm, spanning 302 bands. This covered the entire visual spectral range. The second region was from 753nm to 896nm, or 143 bands, and corresponded to the red edge and NIR shoulder. The latter region was not anticipated to span such a large range, as it is well documented that the NIR quickly plateaus and becomes insensitive to the influence of chlorophyll (Lamb et al., 2002).

These two identified ranges showed that approximately 20 percent of spectral data were associated with chlorophyll. Only 19 bands in the red edge failed to surpass the critical value threshold of ±0.195. An inverted peak near the green reflectance maxima and the red absorption maxima appeared (the latter feature also seen in Figure 3 of Lin et al., 2015), where the correlation coefficient quickly reduced in magnitude before it continued to higher correlation values. This would support the literature which states that, due to saturation, absorption troughs are generally not the most desirable regions for modeling (Le Maire et al., 2004). The correlogram confirmed studies should be concentrated in the visible and red edge regions when chlorophyll estimation is concerned.

The strongest negative associated band was used in the analysis. It was determined through the K-S and S-W test that the distribution of reflectance values at band 702nm were not distributed normally ($p = 0.04$ and 0.01). The reflectance values for

this band varied between 0.08 and 0.48. There was a positive skew of 0.46, suggesting a higher frequency of lower values; this would correspond with more samples exhibiting low reflectance in the red absorption feature transitioning to the NIR. This was expected, as fewer samples saturated at the red well, despite the low threshold for chlorophyll capacity. However, a square root transformation corrected for this skewness, resulting in significant values that passed the K-S and S-W tests (p = 0.074 and 0.118, respectively).

The negative correlation showed that as chlorophyll content increased within the grape leaves, reflectance value decreased; this was in accordance with expectations that the red edge shifts to longer wavelengths for healthy vegetation, which resulted in reflectance reduction as the peak shifts away, while stressed vegetation experience a blue-shift (Ustin et al., 2004). A linear regression of the double axis transformed variables yielded $R^2 = 0.89$, RMSE = ±2.50 µg/g. Four samples skewed the residual plot (sample #s 141, 145, 148 and 158 in Figure 3.9), as they were more than two standard

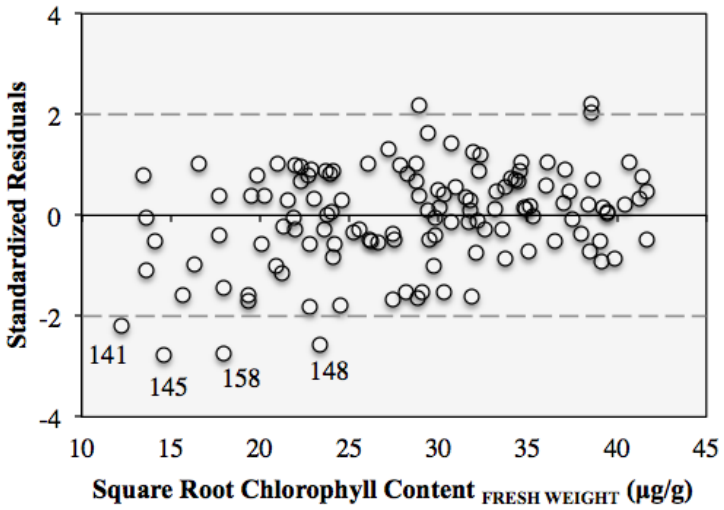

Figure 3.9: Standardized residuals from regression of the single band 702nm with chlorophyll expressed in terms of fresh weight. A double axis transformation corrected for distribution frequencies. Four samples stood out as influential to this model.

deviations from the zero axis; otherwise the residuals were uniform. Given the four samples, the residuals indicated the model was overestimating, as there is more dispersion in the lower half of the scatterplot. These samples were determined as data extremes, rather than outliers, and at an advanced stage of senescence. They were collected on the last sampling day. The four extreme values notwithstanding, the unstandardized residuals had a mean of zero and a normal distribution according to the more lenient K-S test ($p = 0.08$).

The band location and strength of the coefficient of determination complement the findings of Datt (1999) and others, who determined band 710nm as maximally sensitive to chlorophyll content in a eucalyptus study. However, Datt (1999) and Quinn (2010) identified the green peak around 550nm as highly related, whereas this study suggested the next band of interest was on the left shoulder of the red absorption well, around 630nm (Figure 3.8).

3.2.1.4 Datt & Maccioni

Results presented here were generated from the published indices, as well as modified versions. In addition to using the 710nm band as determined by Datt (1999), 702nm was used for this grapevine dataset (as the previous correlogram identified as most correlated to this specific dataset). Consequently, two sets of indices were run (the customized and unaltered published). While the metrics from the original Datt and Maccioni indices were distributed normally with the K-S normality test ($p = 0.20$ and 0.06, respectively), the adjusted index metrics were not, and no transformation altered the distribution to favour normality. This was anticipated, as 702nm was known to be non-normal in distribution. Of the five bands constituting the index equation, only band

710nm was normally distributed (p = 0.097 and 0.099 for the K-S and S-W tests, respectively). Results were presented with this violation acknowledged, and occasioned more stringent assessment of the residuals.

A linear algorithm represented all relationships. Initial results from the SPSS curve estimation function indicated second and third order polynomials would not contribute significantly to a linear equation, and an exponential model would yield a lower coefficient of determination without significantly improving the residual dispersion. Results from the two sets of indices were very similar, but the coefficient of determination, standard error of the estimate, F statistic, and residual scatter plots showed the adjusted Maccioni index to be the best estimate, $R^2 = 0.92$, RMSE ±2.10 μg/g. The R^2 values of the adjusted equations increased the originals by 3% and 2% in their ability to account for variation in chlorophyll (Datt and Maccioni, respectively). Maccioni's index outperformed Datt's in both iterations. However, in every instance, the index variable metrics consistently underestimated values, more noticeably at lower chlorophyll levels.

Compared to the linear relationship established in this study, the final predictive algorithm by Datt (1999) indicated an exponential relationship with chlorophyll, with the index saturating at chlorophyll levels higher than 40 μg/cm^2. Maccioni et al. (2001) also found a linear relationship of

$$\text{Chlorophyll Content } (\mu g/cm^2) = -11.6 + 88.56 * \text{MaccioniIndexValue} \quad (27)$$

when chlorophyll is considered in terms of area. Xue and Yang (2009) noted the Datt Index outperformed the mND Index when considering leafy green crops, while the results of this study found them highly comparable. Main et al. (2011) noted both these indices experienced poorer performance on datasets exhibiting higher chlorophyll content (i.e

savanna trees) than lower pigment containing crops (i.e maize). As the sampling occurring during senescence, a higher chlorophyll range for grapevines was not obtained, which could contribute to the linearity of the model; however, overall the modeling was successful in the range collected.

3.2.1.5 CI red edge

For the chlorophyll index, model tuning was required to identify optimal bands to represent reflectance in the NIR and red edge regions. Band selection significantly impacted the effectiveness of the index as a predictor of chlorophyll, despite it being designed to be conducive to broadband use. The red edge spanned approximately 670nm to 750nm. Rather than averaging the range to obtain a mean position, initially it was hypothesized that the reflectance value at the non-static $R_{red\ edge}$ inflection point (REIP) was the optimal estimate for the red edge index variable. However, the model behaved poorly with this transient reflectance value, yielding a parabolic relationship. A reference point needed to be fixed to observe how samples changed relative to each other at that specific point, rather than how the overall spectral response shifted and migrated. Therefore, the static locations of Gitelson et al. (2003) and Steele (2007) were examined. Steele (2007) proposed any bands within the NIR range and Red Edge range of 755-775 and 710-720nm would be equally efficient, and offered a final selection of 760nm and 715nm. It was anticipated that since that range was developed on a grape leaf dataset, it would be suited to this study. However, the proposed red edge range did not align well; as depicted in Figure 3.10, the lower red edge boundary (710nm) left a segment of under utilized data to the left of it. There were more highly correlated bands at shorter

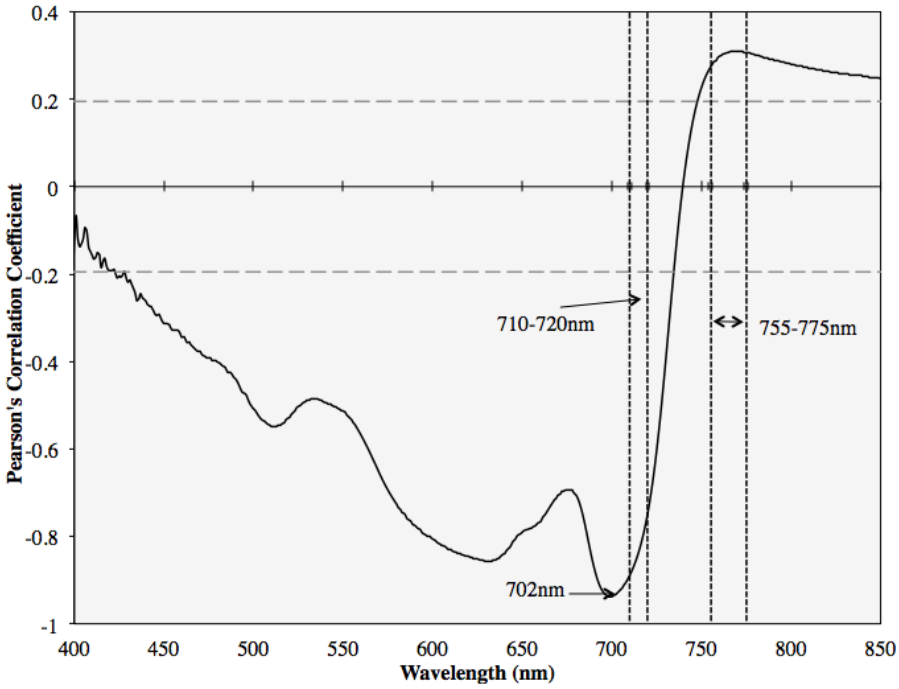

Figure 3.10: The grey vertical dashed lines depict the two spectral regions presented in Steele (2007) that pertained specifically to grapevines and were optimal for the CI red edge index. The correlogram of reflectance to chlorophyll per waveband is included from this study, showing that the first region presented by Steele does not coincide to maximum correlation of this dataset.

wavelengths (the maximum correlation falling at 702nm). This was an unexpected result, and speaks to the range and variation in spectral response observed in grapevine species over their phenological cycle. There was no varietal overlap between these two studies.

Gitelson et al. (2003) reported ranges of 750-800nm for the NIR and 695–740nm for the red edge, and stated all bands within the ranges should perform equally well. These ranges were determined through the identification of continuous extents of RMSE error correlograms that yielded consistently low values when comparing the estimated and measured chlorophyll content. Since the Gitelson et al. (2003) range was more suitable than Steele (2007), the former equation was used, with the amendment that the

correlogram was used to guide the exact band selection locations for the narrowband spectra. This resulted in 768nm and 702 nm as the elected bands.

The index metric passed the K-S test ($p > 0.20$) for frequency distribution once a square root transformation was applied. While Steele et al. (2007) and Gitelson et al. (2003) both had strong linear relationships (e.g. $R^2 = 0.97$), a logarithmic algorithm suited the data better (Figure 3.11). This improved the coefficient of determination (R^2) to 0.92, and achieved a standard error of the estimate of ±2.11 μg/g and residuals with a normal distribution and a mean of zero. The final algorithm was

$$\sqrt{\text{Chlorophyll Content (μg/g)}} = 20.2\ln(x) + 23.85, \quad (28)$$

where *x* is the square root of the CI Red Edge Index.

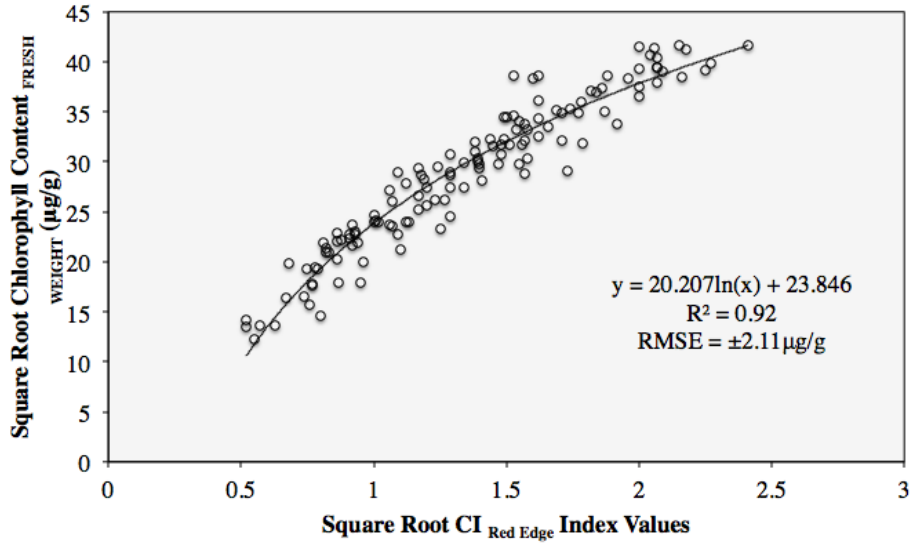

Figure 3.11: A logarithmic relationship between CI Red Edge Index values and chlorophyll content once data was transformed for normality.

3.2.1.6 Modified NDVI

The equation $(R_{750} - R_{705})/(R_{750} + R_{705} - 2R_{445})$ yielded index values that ranged from 0.068 to 0.69, and complimented the range determined by Le Maire et al. (2004). The distribution was multimodal, and as such, failed the S-W test but passed the K-S test

(p = 0.003 and 0.08, respectively). Curve estimates indicated that a quadratic equation was most suitable for modeling. It yielded a coefficient of determination of 0.92 and a standard error of the estimate of ±2.11 µg/g (Figure 3.12). The residuals showed a clustering of three samples that were more than two standard deviations away from the zero axis at the higher pigment contents. However overall, the residuals had a mean of zero and a normal distribution by both the K-S and S-W test (p > 0.05). Comparatively, the results of Sims and Gamon (2002) indicated their model had limitations, notably, their NDVI index values exceeded one. This would require the numerator to be larger than the denominator within index calculations, and most likely occurred in situations where surface scattering was significant enough to cause large reflectance values in the normally suppressed blue region. While results presented here performed better than Girma et al. (2013) and Sims and Gamon (2002) for R^2 values (0.65 and 0.83, respectively), both were investigating more complex scenarios than this study (photosynthetic bark potential and across-species investigation).

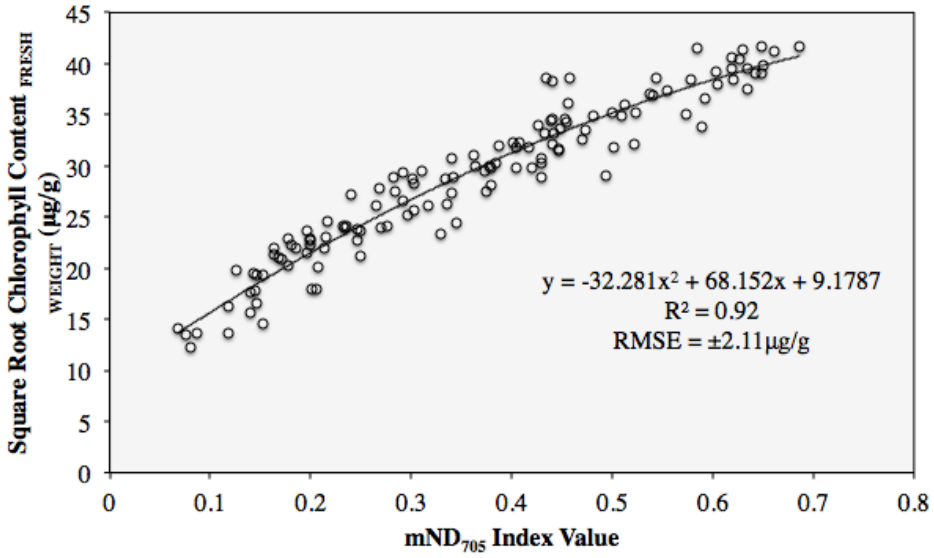

Figure 3.12: A quadratic relationship between the modified NDVI index values and chlorophyll content once data was transformed for normality.

3.2.2 Derivatives

3.2.2.1 Theoretical Expected Results

The general behaviour of the derivative can be characterized by the migration of the red edge peak in relation to chlorophyll content (Gitelson et al., 2003; Rock, Hoshizaki, & Miller, 1988). There is a shift towards the shorter wavelengths (a blue shift) associated with increasing stress and decreasing chlorophyll content (Sanches et al., 2014). This slope change is a result of the decrease in the strength of the absorption in the red region, and the contraction of the feature, as well as changes in the strength of the NIR reflectance (Cho & Skidmore, 2006). There were two important influencing factors in the literature concerning derivative spectra characteristics. The first is the presence of more than one peak in a single derivative reflectance within the red edge region. Multiple local peaks were found in studies by Gitelson et al. (1996) and Zarco-Tejada, Pushnik, Dobrowski, and Ustin (2003), however, they were not observed in this dataset. The multiple peaks resulted from different sampling methodologies that included hemispherical reflectance spectra, suggesting that very fine changes in slope could be captured. This multi-peak is different than the second spectral behavioural characteristic: the double peak. Sometimes there is confusion with terminology in the literature between double-peak and multi-peak features. Zarco-Tejada et al. (2003) referred to two peaks within one spectral reflectance sample as double-peak, but Gitelson et al. (1996) and Le Maire et al. (2004) demonstrated there can be more than two; up to four and six were found, respectively. Therefore, the phenomenon of a single sample having multiple local peaks within the red edge region was described as multi-peak in this paper. The double peak was considered the clustering of two distinct groups of samples in the red edge

region resulting from a chlorophyll threshold, causing derivative peak jump. Reasoning behind this feature is explored in section 4.1.1.2.

3.2.2.2 Summary

The Double Difference Index presented by Le Maire et al. (2004) was the most robust and overall strongest performing metric of the derivative data section of this paper, followed by the single band index with the highest R^2. The other indices were adequate, but with limitations of either insensitivity or lack of range, as indicated by their residuals. MaxD was the poorest performing metric, and a detailed description of its results was not included. All other metric details are summarized in Table 3.2 below.

Table 3.2: Summary of the performance metrics results for the first order derivative index regression models.

Index Name	Coefficient of Determination R^2	Standard Error of the Estimate RMSE	Model Expression
REIP	0.79	±3.40 µg/g	Linear
AltREIP	0.73	±4.50 µg/g	Linear
MaxD	0.61	±4.60 µg/g	Linear
DDI	0.89	±2.45 µg/g	Linear
R'mSR Index	0.87	±2.71 µg/g	Quadratic
R'mND Index	0.86	±2.70 µg/g	Linear
Single Band 690 Index	0.90	±2.40 µg/g	Linear

3.2.2.3 REIP

The wavelength positions of maximum inflection of the derivative reflectance spanned a total of 31nm, from 685 to 716nm. Results from the literature support these findings, for example Curran at al. (1991) with values from 686-724nm for a broadleaf plant Amaranthus, and 690-717nm for Douglas fir needles by Quinn (2010). But some studies, reported positions as far as 740nm for maple leaves (Gitelson et al., 1996).

The REIP frequency distribution failed for both normality tests (K-S and S-W test). It was strongly non-symmetrically bimodal, with the lower wavelengths having a much higher frequency of inflection positions than the higher and longer wavelengths. There was also a noticeable gap between the two modes, from approximately 705-710 nm. Transformations were unable to correct for bimodal distributions. In a theoretical sense, bimodal could indicate the dataset came from two different underlying populations. However, in a comprehensive review of hyperspectral remote sensing of vegetation pigments, Blackburn (2007) suggested wavelength position of the red edge inflection point jumps at a certain chlorophyll concentration threshold, impeding the predictive potential of the chlorophyll-REIP relationship. Consequently, the bimodal pattern is a known issue inherent in some types of derivative analysis, rather than an anomaly with this dataset.

A simple linear regression generated a coefficient of determination of 0.79 with a standard error of the estimate of ±3.40 μg/g. However, an inverted U-shaped pattern in the residuals, with the lower and upper chlorophyll content underestimated, indicated the relationship between the two variables was not linear. While correcting for the residual pattern and increasing the R^2 to 0.90, a second or third order polynomial resulted in inaccurate estimations for chlorophyll content at higher and lower REIP positions (Figure 3.13). Since it was demonstrated in the literature that higher REIPs are a realistic possibility for broadleaf vegetation, it was not advisable to use a quadratic equation as the regression model for grape leaf chlorophyll estimates when considering REIPs. It was determined that due to the bimodal nature limiting the dataset, only a linear regression could be used, as no other curve estimates were appropriate. Overall the REIP as a

chlorophyll estimator is limited due to the nature of the double peak feature.

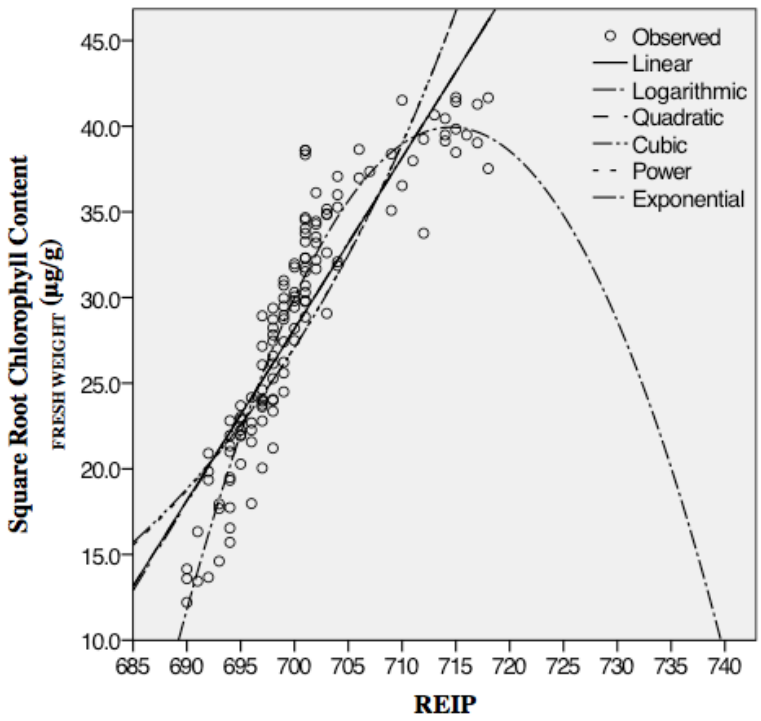

Figure 3.13: Scatterplot of the REIP and Chlorophyll $_{\text{Fresh Weight}}$ depicting the unsuitability of some of the function equations for REIP positions outside the given chlorophyll range of this dataset, particularly for leaves with higher chlorophyll content.

3.2.2.4 AltREIP: Linear Interpolation REIP Index

The application of the linearly interpolated index as an alternative for the REIP was unsuccessful in the case of this senesced grapevine dataset, due to the inability to correct for residual patterns. The linear regression coefficient of determination was 0.73 with a RMSE = ±4.50 µg/g. The model had limits despite avoiding the peak jump in the data (Blackburn, 2007). An exponential or second order polynomial would increase the R^2 = 0.733 and 0.84, respectively, but both models retained curved patterns in the residuals, and a quadratic equation would have vastly over- and under- predicted values outside the data range. Overall this method was not a better estimate of chlorophyll over the more traditional REIP method, nor did it perform as well as the baseline single

derivative model. Comparative studies such as Girma et al. (2013) for bark chlorophyll estimation had similar difficulties with their linear extrapolation method, $R^2 = 0.65$.

3.2.2.5 DD Index (DDI)

No combination of scaling and transformation normalized the frequency distribution of the DDI metric sample values. While initially it was speculated that the DDI used the derivative spectra in the equation calculations, the best results of the DDI were obtained when the derivative reflectance were used in the selection of band locations, and the unaltered reflectance was used within the index calculations. In determining the four band locations, R_b, $R_{b+\Delta}$, R_a and $R_{a+\Delta}$, three separate methods were considered. Firstly, Le Maire et al. (2004) proposed a computational iterative process of band combinations that minimized RMSE error. Secondly, this study aimed to identify key bands via correlogram as a method of model tuning the band selection specifically for grape leaves. Bands 675nm, 702nm, 735nm and 762nm generated the strongest correlation coefficients between derivative reflectance and chlorophyll. However, these bands were unsuccessful in creating a stronger model. Ultimately, a refined band selecting of 670nm, 700nm, 720nm, and 750nm yielded the best results of these three methods (Haboudane et al., 2008).

The DDI resulted in a strong positive relationship, with values having a strong positive skew frequency distribution, in addition to spanning both negative and positive values from -0.22 to 0.13. The data range was comparable to the results -0.20 to 0.15 (Le Maire et al., 2004), but the dispersion differed, as more than half the data points in the DDI were negative for this study, while less than a quarter were for the other. A slight general underestimation of points across the entire chlorophyll range was observed with a linear

regression, and mirrored in the residual pattern. It was confirmed that these points were not a result of the peak jump, where one peak group is represented in the lower estimates. As the vast majority of the samples fell into the 1st peak group, the 2nd peak group were characteristic of very high chlorophyll content, which the residuals showed as well characterized. The DDI index performed well, and sufficiently accounted for the double peak feature in the dataset.

The final coefficient of determination was 0.89 with a RMSE = ±2.45 μg/g. This was an improvement of the independent variable's (DDI) ability to account for the variation of the dependent variable (chlorophyll), as compared to the same index but with bands determined by Le Maire et al. (2004) and this study. The final linear regression equation was determined as

$$\sqrt{\text{Chlorophyll Content (μg/g)}} = (7.38\ (\text{DDI value}) + 3.17). \quad (29)$$

3.2.2.6 R'mSR Index

For the R'mSR index, a second order polynomial equation was the most appropriate. The quadratic equation was determined to be

$$\sqrt{\text{Chlorophyll Content (μg/g)}} = 1.80 + 2.89 * \text{R'mSRindexValue} - 0.88 * \text{R'mSRindexValue}^2 \quad (30)$$

for a coefficient of determination of 0.869, with a RMSE = ±2.71 μg/g (Figure 3.14). The second variable in the model equation significantly added to the improvement of the linear model by an R^2 change of 3%. However, there was still a slight dispersion of the residuals at very low chlorophyll levels, resulting in non-normality, which suggested that index would have predictive limitations for the lower foliar pigment contents, and may not be appropriate for senescing datasets. It was noted that the experimental dataset used

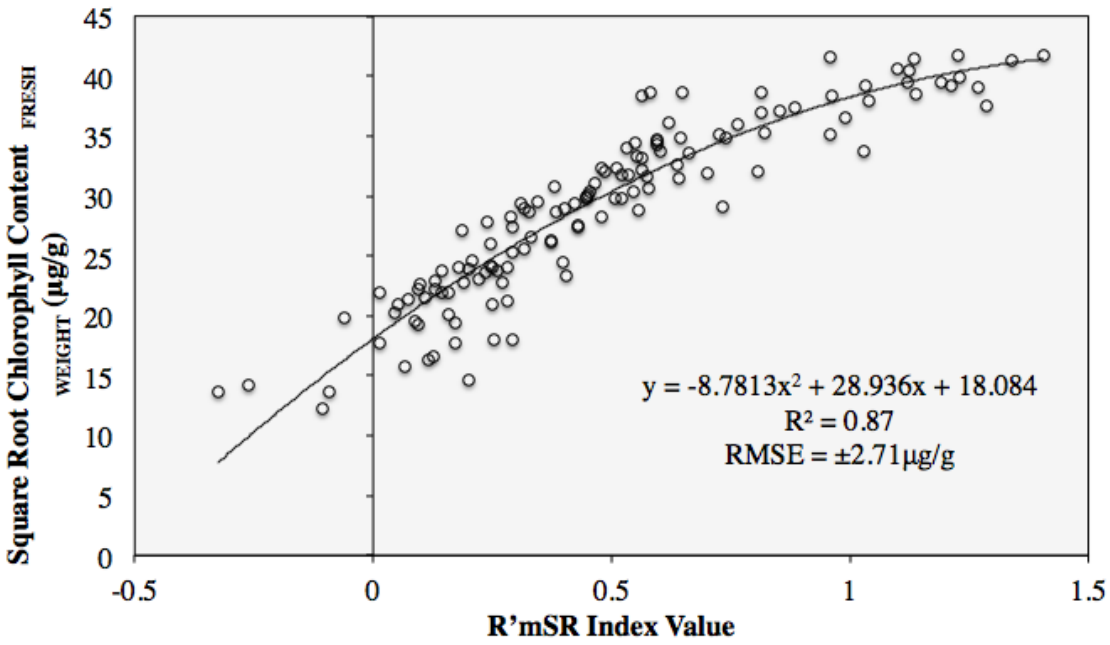

Figure 3.14: Results from the modified Simple Ratio Index with derivative reflectance spectra in a quadratic model. Heteroscedasticitiy was noted at the lower end of chlorophyll content levels**.**

to develop this index lacked very low chlorophyll content, and the index generated few negative values, all of which were small (> -0.1), the majority falling within a range of 0 to 3 (Le Maire et al., 2004). This model operated well at the higher pigment levels, but due to the nature of the quadratic relationship, the index's ability to detect changes of high chlorophyll will saturate; index values would continue to increase with very little change in chlorophyll.

3.2.2.7 R'mND Index

The R'mND index was best represented with a linear function, but the study by Xue and Yang (2009) determined it as nonlinear, with lettuce leaves providing low chlorophyll measurements, revealing a dramatic drop in index values (Figure 3.15). While the dataset in this study alludes to the beginning of the curve captured in Xue and Yang (2009), there insufficient justification statistically for the use of either an

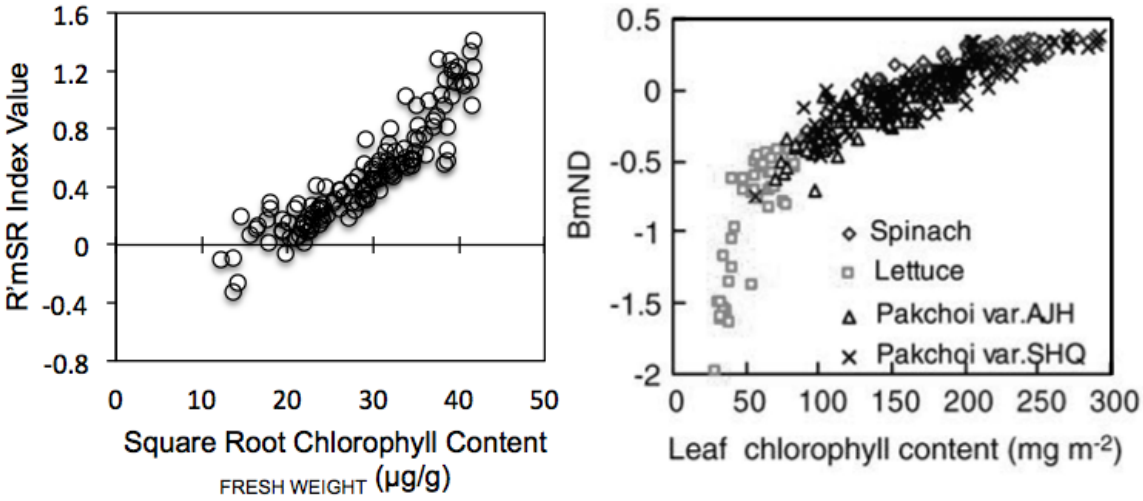

Figure 3.15: Comparison of the results between this study, left, and Xue et al. (2009), right. While this study presented a linear model, the lower range of chlorophyll values indicate that start of a curve, which was fully characterized in the other study. Differences in model functions were attributed to variable transformations and chlorophyll expressed in different units.

exponential or quadratic equation (the latter improving the model only by 1% for the R^2 and did little to correct for the pattern in the residual error). The coefficient of determination was 0.87 with a RMSE = ±2.70 µg/g. Residuals had a mean of zero, and were distributed normally according to the K-S but not the S-W test significance values. This was due possibly to the leptokurtic nature of the frequency distribution, and the influence of three very low chlorophyll values.

The index developed by Sims and Gamon (2002) addressed indices that correlated poorly with chlorophyll when faced with multiple species of vegetation. Despite having several varietals in this study, it appeared the grape leaves collected lacked the extreme low chlorophyll levels required to capture this curve. Therefore this result is presented with the caveat that the full relationship may not have been characterized. It also suggests that the index may not be suitable for vegetation exhibiting high chlorophyll levels, as it will saturate quickly.

3.2.2.8 Single Band (correlogram)

This method investigated the relationship of a single derivative reflectance value at a specific wavelength position to chlorophyll. The correlogram of the derivative data indicated band 721nm was the strongest positive Pearson's correlation coefficient (0.946), and band 690nm had the strongest negative correlation coefficient (-0.948). Of the 400 bands analyzed (from 400nm to 800nm), only 77 fell below the critical value threshold of ±0.195, and confirmed this region of the spectrum is influenced highly by chlorophyll. These two positions corresponded to approximately either side of the REIP. The start of the steep red edge was near 690nm, and is less susceptible to pigment saturation, while 721nm is around the top of the red edge, on the shoulder of the NIR plateau.

Neither of the bands were normally distributed, with 690nm being platykurtic and 721nm being multimodal. Both of these features are very difficult to correct with transformations. The derivative bands produced strong linear relationships with chlorophyll, however there were distinct patterns in the residuals. 690nm had the slightly superior residual scatterplot with six points falling just slightly outside two standard deviations, while 721nm had four beyond three standard deviations from zero. Band

690nm had an F-stat value of 1138 compared to the 1084 for 721nm, and a smaller residual sum of squares of 7.3 versus 7.7. It was decided that 690nm was marginally more suitable, with an $R^2 = 0.90$ and RMSE = ±2.40 μg/g, and was regressed to generate the equation

$$\sqrt{\text{Chlorophyll Content (μg/g)}} = (-181.7 * \text{DerivativeBand690nm} + 4.1). \quad (31)$$

When comparing these results to Le Maire et al. (2004), 742nm was their most correlated single band. The values of this index (the value of the derivative reflectance) were similar, with neither appearing to go above 0.005. Band 690nm was also determined to be the strongest first derivative band correlated with algae chlorophyll in an aquatic study, suggesting that the fundamental relationship is strong enough to be spectrally preserved through water turbidity (Luoheng & Rundquist, 1997).

3.2.3 Continuum Removal

3.2.3.1 Theoretical Expected Results

As established in the literature, there is an expected spectral response to changes in chlorophyll (Quinn, 2010; Sanches & Filho, 2014). These changes are attributed to different factors, such as nutrient deficiencies, infestations, and, in this study, senescence. This degradation of chlorophyll affected the visible region from 550nm to 730nm, where there is a strong absorption of red light by this pigment within leaves (Sims & Gamon, 2002). Both the left and right areas under the continuum will contract toward the max band centre with the decrease in chlorophyll, and broaden with an increase in foliar pigment (Figure 3.16). The area differences are preserved in the horizontal and vertical shifts and total area under the curve, though subtle changes become more significant (Kokaly & Clark, 1999). It was expected that the left area consistently would be larger than the right area. Due to the contact probe and leaf clip attachments used with the ASD, there was very little external light scattering. The spectra had very few confounding external variables, and as a result, captured the leaf reflectance thoroughly.

3.2.3.2 Summary

The results below indicate that within the continuum removed metrics generated, the Total Area metric proved most suitable as a chlorophyll estimator. Despite initial concerns that transformations of the independent axis could not normalize the frequency distribution, the Total Area had a final coefficient of determination of 0.93. In addition to

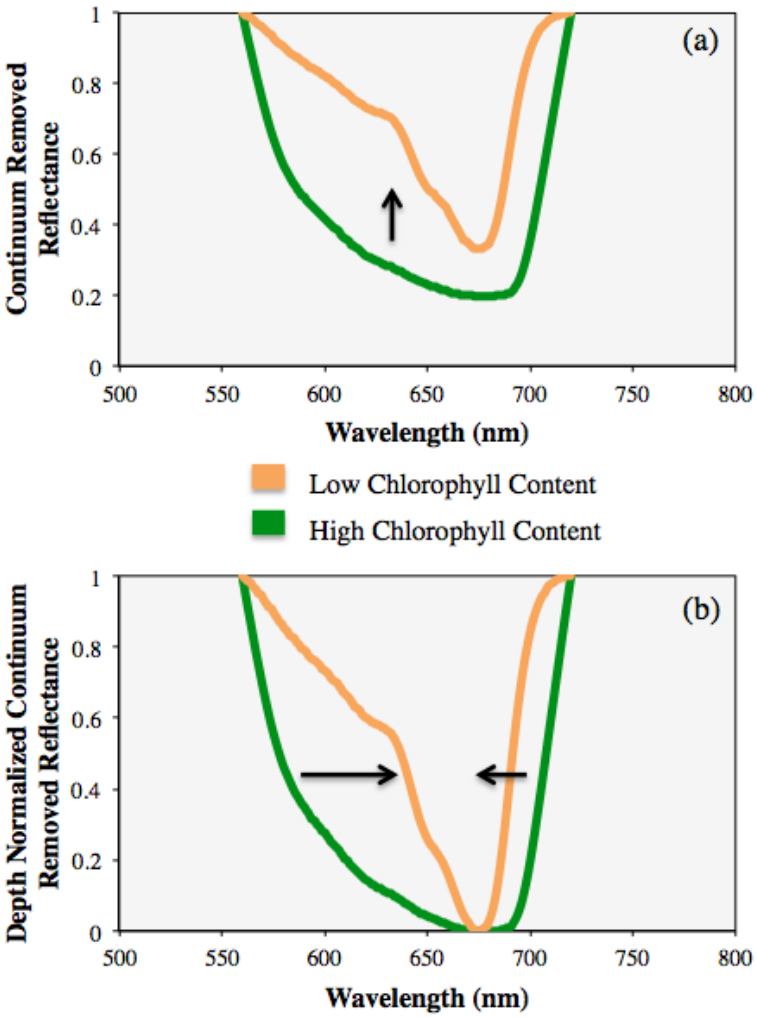

Figure 3.16: Demonstrating the changes in the shape of the continuum removed spectrum with changes in chlorophyll content over the red absorption feature. (a) The max band depth migrates upwards as the absorption decreases and (b) the contraction of the feature is seen with the depth normalized spectra. Migration occurs from healthy (green line) to senescent leaves (yellow line) based on chlorophyll content changes.

strong predictive power, it also had the highest explanatory power for the equation variables ('goodness of fit'). Other metrics proved unsuitable for a variety of reasons including heteroscedasticity and poor variable correlation. The TA metric was selected for the validation method to determine an optimal foliar biochemical estimation equation for grape leaves.

Results determined to be non-influential, and that did not contribute significantly to results, were not reported here. Metrics not selected were the Maximum Band Depth Location, which is the same as the Band Symmetry point between the Left and Right Areas, and the value of the Maximum Band Depth. While some studies had success with Maximum Band Depth, they often utilized different statistical methods or were concerned with other biochemical constituents (Curran et al., 2001; Kokaly & Clark, 1999; Mutanga & Skidmore, 2003). Girma et al. (2013) had similar results with their asymmetry ratio for bark chlorophyll estimation ($R^2 = 0.05$). Van der Meer (2004) reported more success, but for a geological, not biochemical, study.

3.2.3.3 Left Area

The values generated for the Continuum Removed Left Area (CRLA) metric were not normally distributed (K-S and S-W test significance ≤ 0.001), and no applied transformation corrected this. Results are presented with this violation acknowledged, and a more stringent assessment of the residuals resulted. A scatterplot between the continuum metric and chlorophyll content revealed a curve in the data with a relatively strong, positive relationship. Curve model estimates suggested exponential and quadratic functions had the most uniform variance in the residuals scatterplot, with the exponential spread being slightly more centered around the mean of 0. The residual sum of squares

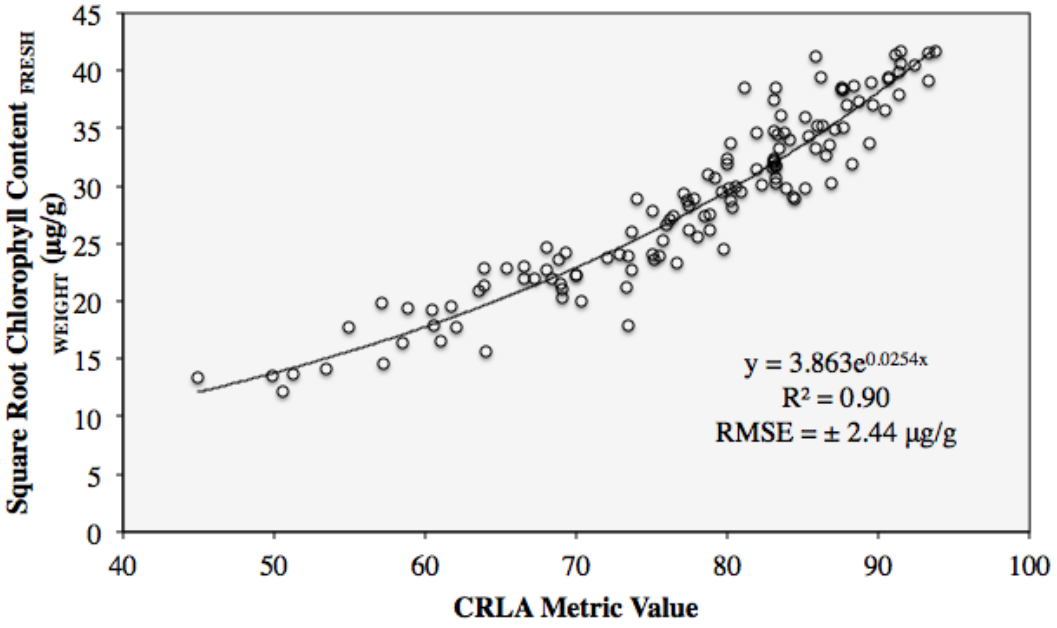

Figure 3.17: An exponential function was determined for the CRLA-pigment regression model.

was smaller for the exponential than for the quadratic, implying a larger variance between observed and estimated values for the quadratic. This translated into larger model error, rendering the quadratic model suboptimal for the data, despite having very similar coefficients of determination and other statistical results (Field, 2009). The exponential model was determined the best estimate, and a non-linear regression was performed, R^2 = 0.90, RMSE = ± 2.44 µg/g (Figure 3.17). The residuals were normally distributed, with both K-S and S-W tests having a significance of ≥ 0.195, and a mean of 0.0001. The R^2 was the proportion of the variation in the values of the square root of chlorophyll concentration explained by the least square exponential regression line of chlorophyll on the continuum removed left-of-max-band-depth accumulative area of reflectance.

3.2.3.4 Right Area

The data range for the accumulative area under the continuum removed line, from the maximum band depth to the absorption feature's right limit, was small compared to

the left area, from 16.87 to 33.46. The Continuum Removed Right Area (CRRA) was one of the few metrics normally distributed, confirmed through the K-S and S-W tests (both $p \geq 0.05$). It was a positive linear relationship between the input variables, as the area under the continuum increased with chlorophyll content. This was consistent with the predicted findings, since the right side of the continuum encompassed the red edge, and a spectral shift towards the right in response to increased chlorophyll would increase the overall area under the continuum curve for the right side. Although both input variables were distributed normally, a scatterplot revealed a tendency towards increased dispersion when both variable values increased. Although the Right Area had a normal distribution frequency, and an apparent linear relationship with chlorophyll, a heteroscedastic pattern in the residuals caused the variance to increase across the residuals, making the model insubstantial and unsuitable for chlorophyll estimation.

3.2.3.5 Total Area

The data range for the cumulative total area under the continuum removed line for the complete absorption feature was from 60.88 to 119.40. The frequency distribution of the Continuum Removed Total Area (CRTA) metric was not distributed normally ($p \leq 0.001$ for both K-S and S-W test), due to the skew of the left area. No applied transformation corrected for skewness, and results were presented with this violation acknowledged, resulting in a more stringent assessment of the residuals. A scatterplot between the continuum metric and chlorophyll content revealed a strong positive linear relationship between the input variables. A curve estimate in SPSS identified linear, quadratic, cubic, power, and exponential equations as possible model candidates. A simple ordinary least squares linear regression was performed, $R^2 = 0.90$, RMSE = ±2.33

μg/g. However, a pattern in the residuals was prominent enough to warrant consideration of the nonlinear alternatives. The higher chlorophyll extrapolations for the models were all very similar; the difference between the models was in the lower values and the possible extrapolations beyond the range of the dataset. The quadratic was the simplest model that explained the most variation in foliar content by the reflectance metric, and still presented feasible values for the lower extrapolated chlorophyll values. However, it was noted that due to the senescence, the leaves would be unable to advance much further into the chlorotic stage before leaf abscission would occur. The quadratic modeling yielded an equation with variable coefficients and intercept determined to be

$$\sqrt{\text{Chlorophyll Content (μg/g)}} = 3.07 - 0.064 * x + 0.001 * x^2, \quad (32)$$

where x is the CRTA metric value. The results were a coefficient of determination of 0.93 and RMSE = ±1.96 μg/g. The non-standardized residuals were distributed normally (confirmed with both K-S and S-W test significance ≥ 0.20), and achieved a mean of zero. This quadratic function was determined to be the optimal results attainable for this metric.

3.2.4 Validation

The continuum removed depth normalized total area metric was selected for the reflectance validation process. This provided an additional opportunity to assess and characterize the differences when reflectance measurements are used in conjunction with two discrete chlorophyll units, as reflectance analysis focused solely on the fresh weight chlorophyll values in the previous sections. Therefore, the calibration and validation was performed on chlorophyll expressed both in terms of unit fresh weight, and unit surface area, with the CRTA metric values. As the data was split into two subsets, the results

Table 3.3: Summary of the performance metrics results for the reflectance validation methods.

Validation Method	Chlorophyll per unit area (µg/cm$^{2)}$		Chlorophyll per unit fresh weight (µg/g)	
	Coefficient of Determination R^2	Standard Error of the Estimate RMSE	Coefficient of DeterminationR^2	Standard Error of the Estimate RMSE
Data Splitting	.92	±2.29 µg/cm^2	0.93	±111.01 µg/g
Jackknifing	.92	±2.38 µg/cm^2	0.93	±109.05 µg/g

yielded different coefficient of determination than obtained previously in the continuum removed section (3.2.3.5).

3.2.4.1 Summary

The results of the data-splitting and jackknifing validation are shown below (Table 3.3). It was expected that the jackknifing would have lower coefficients of determination and higher standard errors of the estimate, as it is a more robust model. However, the model error was actually smaller for the jackknifed results when considering chlorophyll in terms of fresh weight. This suggests that the jackknifing validation method was more suitable than the data-splitting method for the reflectance data processing. The model errors remained consistent over both methods, as well as error dispersion tendencies.

3.2.4.2 Calibration

The reflectance data was split in an identical fashion as the SPAD calibration and validation datasets. Testing for normality indicated that the CRTA calibration dataset of 65 samples did not obtain a normal distribution, regardless of transformation attempts. The chlorophyll subsets for both area and fresh weight expression were distributed

normally, and therefore no transformation was needed. The curve estimate feature in SPSS indicated that quadratic, cubic, power and exponential models were all suitable functions, and therefore the residuals were used to determine the optimal model. The additional variable of the third order polynomial added no significant explanatory power to the model, and was eliminated. The inverted apex of the second order polynomial negatively impacted the lower range of CRTA values by overestimating them (the quadratic curvature veering away from the points) and therefore was eliminated. The power and exponential functions were both very similar, but as the power function intersects with the X-axis and could yield negative chlorophyll values, the exponential equation was the final selection. The models yielded a coefficient of determination of 0.93 and 0.92 for the chlorophyll expressed as fresh weight and as area, respectively. The standard errors of the model estimates were ±112.72 μg/g and ±2.24 μg/cm^2. The residuals were normally distributed and had a mean of zero. No patterns or anomalies existed in their dispersion. The calibration equations were

$$\text{Chlorophyll Content } (\mu g/cm^2) = 0.101 * \exp^{(0.048 * CRTA)} \quad (33)$$

$$\text{Chlorophyll Content } (\mu g/g) = 9.989 * \exp^{(0.043 * CRTA)}, \quad (34)$$

where CRTA is the continuum removed, depth normalized, total area under the continuum metric.

3.2.4.3 Validation

Using the reflectance metric in the calibration model, two estimated chlorophyll datasets were determined. These datasets were compared with the measured chlorophyll content by a goodness of fit linear regression, with the estimated chlorophyll as the dependent variable. The following relationships were obtained:

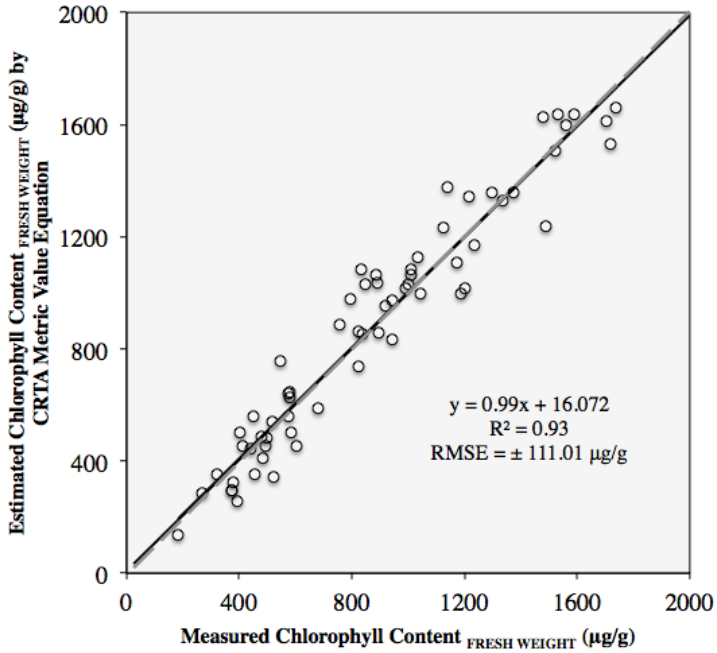

Figure 3.18: Results of the data-splitting validation of the highest performing reflectance metric, CRTA. Comparison of the measured chlorophyll expressed in terms of fresh leaf weight with the estimated chlorophyll determined using the calibration model.

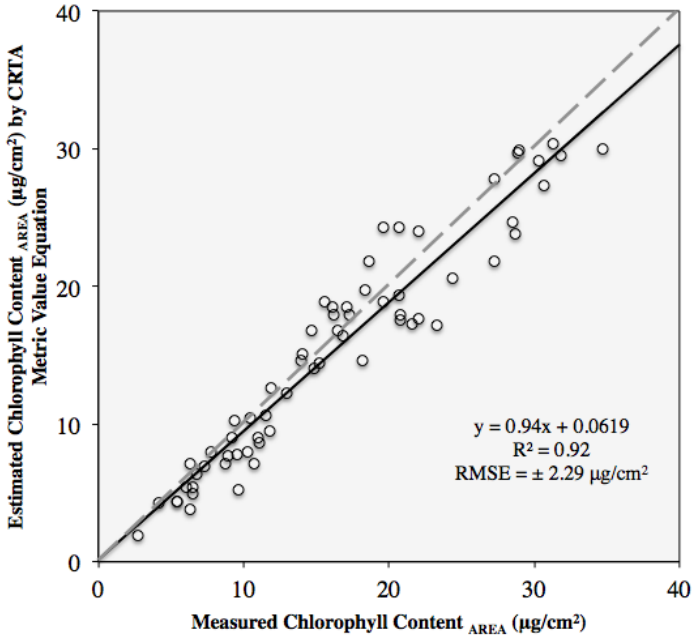

Figure 3.19: Results of the data-splitting validation of the highest performing reflectance metric, CRTA. Comparison of the measured chlorophyll expressed in terms of leaf surface area with the estimated chlorophyll determined using the calibration model.

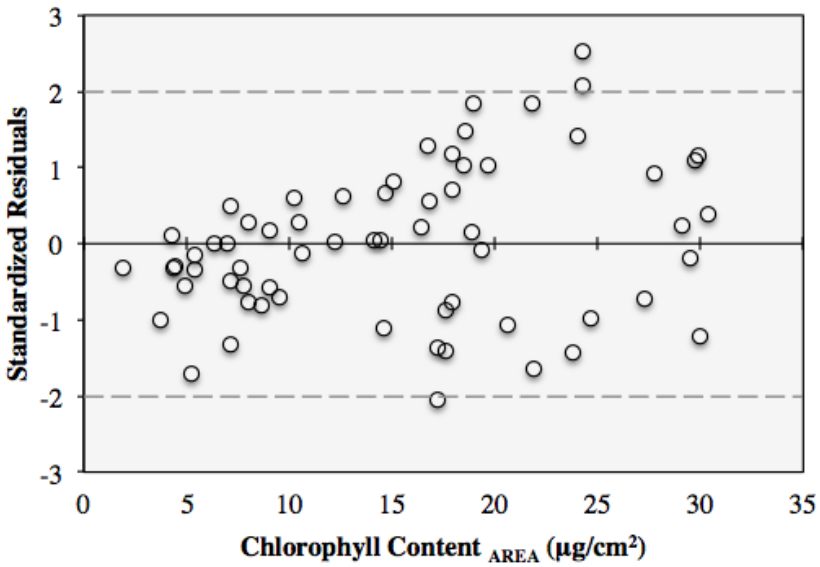

Figure 3.20: Results of the validation of the calibration equation for chlorophyll estimates derived from CRTA equations.

The coefficients of determination were 0.92 and 0.93 for chlorophyll in terms of area and fresh weight, with corresponding RMSE's of ±2.29 µg/cm^2 and ±111.01 µg/g (Figure 3.18 and 3.19). This validated the calibration model as strong and accurate, with the inference that the calibration model had the potential to be effective in estimations beyond the dataset from which it was created. That said, the model with chlorophyll is expressed in terms of area had heteroscedastic residuals (Figure 3.20). Comparatively, the fresh weight residuals fared better, as they had constant variance, therefore qualifying it as the more appropriate model expression.

3.2.4.4 Jackknifing

This resampling approach was performed to further the modeling analysis and assessment, while reducing bias from the dataset (Lawrence & Labus, 2003). The overall accuracy for the continuum removed depth normalized total area reflectance metric was R^2 = 0.92 for chlorophyll per unit area with a RMSE of ±2.38 µg/cm^2 for chlorophyll ranging from 2.21 to 35.82 µg/cm^2, matching well to the measured data range. The other

model obtained a coefficient of determination of 0.93 for chlorophyll per unit of fresh weight, with a RMSE of ±109.05 µg/g for chlorophyll ranging from 148.94 to 1736.89 µg/g. These jackknifed regression results are less sample-dependent than the previous data-split regression, and better reflect the grapevine population of interest (Meeuwig & Peters, 1996). The regression equations for the jackknife validation were

$$\text{Chlorophyll Content}_{\text{predicted}}\ (\mu g/cm^2) = 1.018 * \text{Chl}_{\text{measured}} - 0.92, \text{ and} \quad (35)$$

$$\text{Chlorophyll Content}_{\text{predicted}}\ (\mu g/g) = 0.916 * \text{Chl}_{\text{measured}} + 68.42 \quad (36)$$

for fresh weight measurements and area, respectively, as seen in Figures 3.22a and b.

As with the validation of the data-splitting, the jackknifing also showed heteroscedasticity in the residuals of the chlorophyll expressed in terms of leaf surface area, with a slight increase in underestimation at the higher values (as indicated by the positive residuals in Figure 3.21). The residuals of chlorophyll in terms of fresh weight did not demonstrate this pattern.

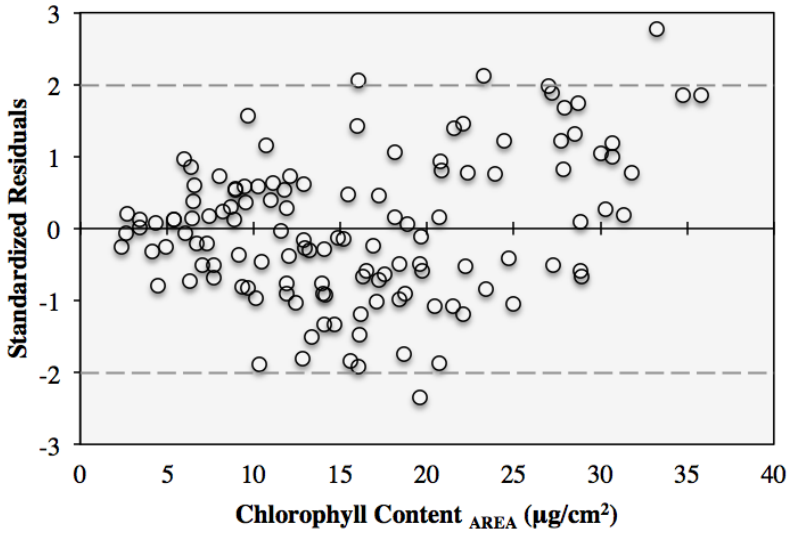

Figure 3.21: Residual results of modeling estimated chlorophyll expressed per unit of leaf area (through jackknifing method of nonlinear CRTA metric model) and destructively determined chlorophyll content.

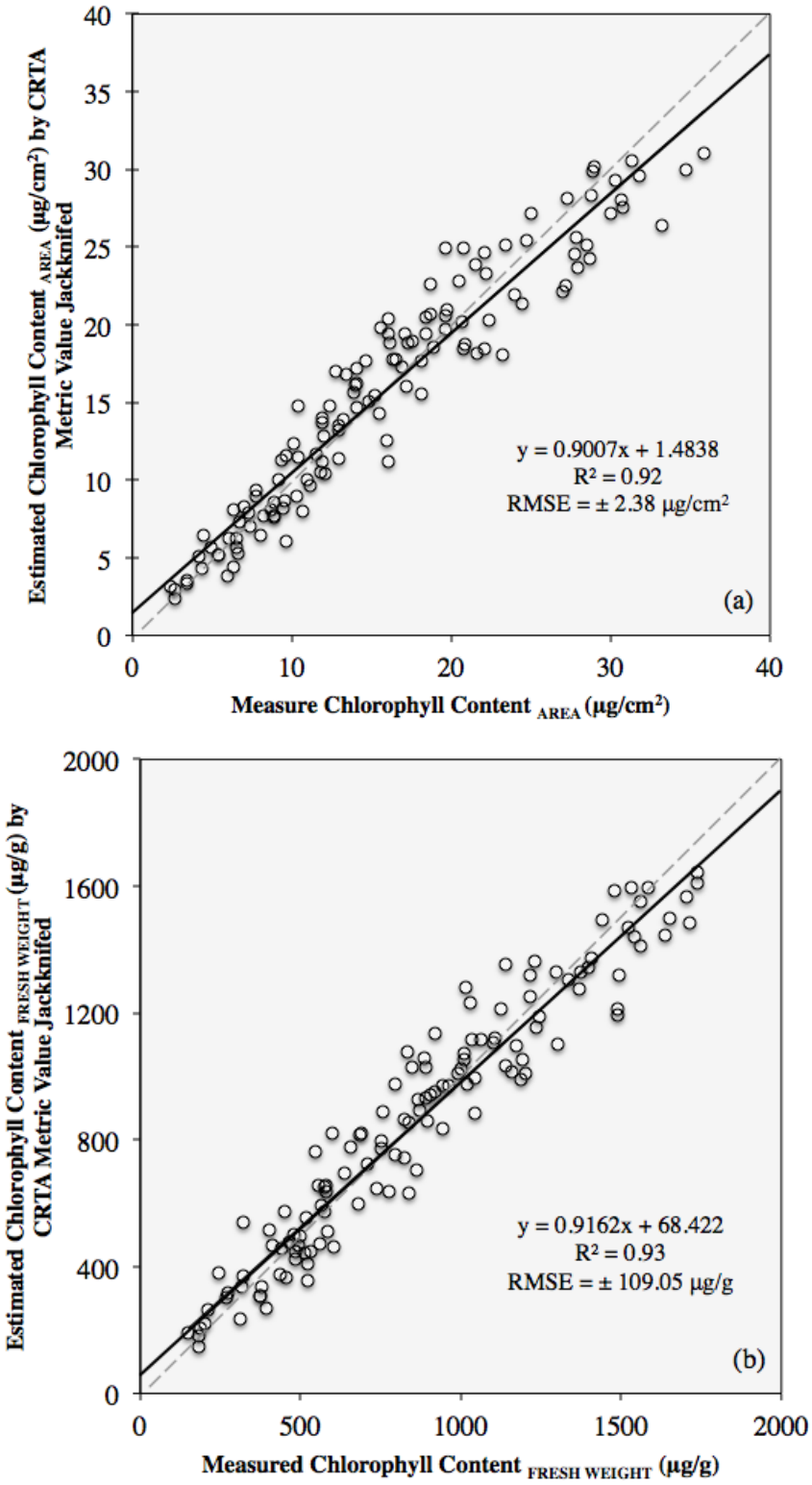

Figure 3.22: Jackknife validation of the highest performing reflectance metric, CRTA. (a) Chlorophyll expressed in terms of leaf surface area. (b) Chlorophyll expressed in terms of fresh leaf weight.

The high coefficient of determination and constant variance of the residuals when chlorophyll was expressed as fresh weight indicated it was the model with more robustness and good predictive power, and suitable for chlorophyll estimation in grapevine leaves when using reflectance measurements.

Chapter 4 Discussion

4.1 Data Specific

4.1.1 Reflectance

4.1.1.1 Baseline & Indices

This section examines the overall performance of reflectance indices compared to the continuum and derivative results. Three findings are discussed; the mND_{705} index as the best reflectance metric, improved modeling from waveband locations offset from the maximum absorption feature, and an unexpected outcome from the correlogram between NIR and pigment content correlation. Lastly, the role of the sampling strategy in the success of these reflectance models is investigated.

Reflectance indices were used to determine if an increased spectral sensor resolution and a hyperspectral processing approach could better capture the spectral-pigment relationship than the broadband counterpart. While hyperspectral sensors capture fine spectral subtleties that characterize minute influences from foliar biochemicals, Stagakis, Markos, Sykioti, and Kyparissis (2010) noted that complex and modified spectral indices do not appear more accurate than their simpler, broadband, predecessors. Based on the results in this study from the four reflectance indices selected for detailed analysis, it was concluded that the indices serve chlorophyll estimation well, despite suffering from limitations and discrepancies when applied quantitatively. Broge and Leblanc (2000) mirrored this finding; a simulated study found narrow-band ratios pertaining to the red edge did not necessarily predict vegetation characteristics better than the classic multispectral models. Furthermore, in this study reflectance indices performed better overall at predicting chlorophyll than derivative indices presented (R^2 = 0.61 to

0.90). However, the same cannot be said for the continuum removal approach, as the total area under the continuum was the highest performing metric in this study.

The $Maccioni_{adjusted}$, mND_{705}, and the $CI_{Red\ Edge}$ index generated equivalent performance metrics, $R^2 = 0.92$. Of these, the mND_{705} index was ultimately determined to be the most suitable model for chlorophyll estimation. Indices that were specifically selected for their robustness to vegetation variations still required extensive model tuning. The mND_{705} index did not require data-optimized band selection. This suggests an additional robustness in the model's predictive abilities.

Indices using wavelength positions that were off-peak from the maximum absorption feature generated better results (Main et al., 2011). This was repeatedly stated within the literature, and occurs because off-centre bands are more sensitive to variations in chlorophyll, especially in higher concentrations, as they do not saturate at lower thresholds (Girma et al., 2013; Sims & Gamon, 2002). Additionally, because reflectance in the red region does not decrease linearly with pigment concentration, the predictive capabilities are limited in index ratios due to eventual model insensitivity (Murphy et al., 2005). The relationship between reflectance in the visible spectrum and chlorophyll content is characterized essentially as nonlinear in nature (Gitelson et al., 2003). Results were congruent with that, as very few regions were linearly related to chlorophyll; even band 702nm (highest correlation) required a square root transformation to model against chlorophyll. As such, it proved critical to employ bands near the red edge for optimal chlorophyll index design (Stagakis et al., 2010).

The Maccioni index used three regions of the chlorophyll absorption feature to fully develop the spectral response, including the maximum absorption feature at 680nm.

It produced a linear model, which is advantageous, as the dependent variable values will not become insensitive to changes in the independent variable. However, all iterations of both the Maccinoni and the Datt indices consistently underestimated the lower chlorophyll values; this suggested a curve in the data spread that was not fully captured at the lower range of values. Additionally, Le Maire et al. (2004) reported poorer model performance with higher chlorophyll content. The saturation of the 680nm may have contributed to these results. Therefore, while the results presented here were successful, any future model performance is expected to be dependent on data range.

The main limitation of the baseline single band index was in its determination. A larger proportion of the NIR was correlated significantly to pigments, up to the 800nm range (Figure 4.1). This was an unexpected result as many vegetation indices utilize NIR bands specifically for their insensitivity (and thus low correlation) to photosynethetically active biochemicals (Murphy et al., 2005). These results ran counter to the theoretical principle that the NIR is not influenced by pigments, but by water content and internal cellular structure. They suggest that a truly unaffected band would be found further along the NIR plateau than previously proposed in the literature (Maccioni et al., 2001; Sims & Gamon, 2002). A Spearman's correlation coefficient could be used to compare variations against the Pearson's correlation, to determine if this NIR range is the result of the statistical metric used. Additionally, a Cook's distance test was used to assess any leverage of specific samples on the metrics potential (Ranjit, 2006). Regardless, the location of the strongest correlation remained within ±5 wavelengths, and the high performance of the indices in this study did not suggest a breakdown of the underlying

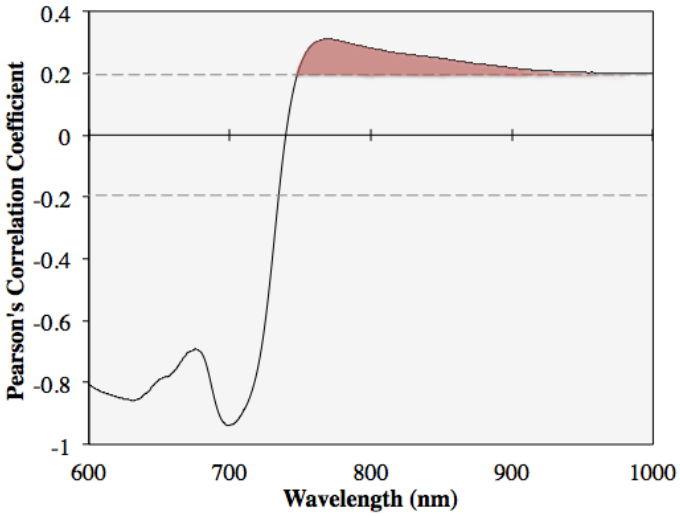

Figure 4.1: The (red) area with significant correlation values extended farther into the NIR than anticipated. This was thought to have been a combination of the Pearson correlation coefficient determination and influential samplings.

principle of NIR insensitivity, as they appeared to have operated as intended. Due to this limitation, however, single band indices are suggested with caution.

The sampling design of this study contributed to the success of these results, in that it investigated the spectral response of vegetation in a very controlled environment. Leaf-level studies inherently control for variations in leaf surface and internal structure because the sampling environment is controlled (Datt et al., 1999). Thus, modeling calibrations report smaller errors than would be observed in the natural environment. This could be seen as a limitation in the sampling design, as it did not fully accommodate and account for the composite spectral response obtained remotely. This is acknowledged, but countered with a reiteration of the study objective: to establish and characterize the relationship between chlorophyll and spectral reflectance obtained from ground validation tools. This did not require all environmental influences to be considered. That said, the degree of success of the models did not correspond necessarily to the degree of correction applied in the models. While Datt & Maccioni, CI red edge, and mND705

indices all performed similarly, they differed in the extent to which they corrected for factors such as surface reflectance, full spectral characterization of pigment absorption, influences of other pigments on spectral response, and backscattering. For example, the use of the 445nm was not contributory to the model, as the blue region did not suffer additive effects from surface scattering. Results are presented in this study with the qualification that they were optimized for the leaf sampling level, and any results would have to be re-assessed for different scales. This important sentiment is shared within the research community (Le Maire et al., 2004).

4.1.1.2 Derivatives

The derivative metrics did not perform as well as anticipated, as none technically achieved performance metrics better than the baseline single-band index. That said, the DDI and single band indices performed comparatively well (R^2 = 0.89 and 0.90, respectively). This section covers the extent of the affect the peak jump had on the derivative results, the performance of metrics that utilize multiple derivative locations in their index, the relationship of the derivative and the REIP, and comments on a possible limitation in derivative analyses within the literature.

There are distinctions that make the DDI more appropriate than the baseline single-band index. The vulnerability of using single bands centres around its specificity, and is limited by (i) the dataset range (ii) the species of leaf (iii) the sampling scale (e.g. leaf-level) and (iv) the control of ambient light scattering by the contact probe. Sample collection was conducted in a laboratory environment with optimal conditions. However, field collection will complicate the spectral signature, obscuring individual band pigment responses. The danger in relying upon one band to represent any one characteristic is

whether the band is expressive of any other confounding variable influencing the reflectance. Therefore it would be desirable to select an index that utilizes more than one reflectance characteristic to quantify a more comprehensive understanding of the dynamic relationship between spectra and chlorophyll. DDI was determined as the best derivative metric for chlorophyll estimation.

The occurrence of the double-peak discontinuity feature was unexpected. In a comprehensive review on technologies and methodologies of hyperspectral remote sensing for pigment estimation, Blackburn (2007) cited a study by Zarco-Tejada et al. (2003) that presented fluorescence as a possible reason for this feature. Although, this better explains multiple local peaks within one reflectance spectrum rather than the bimodal distribution of the REIP position distribution frequency for a dataset. Double-peak discontinuity is more likely the result of natural conditions within the leaves.

Le Maire et al. (2004) determined the peak jump occurred after a chlorophyll threshold of 45 $\mu g/cm^2$ was surpassed. The current study observed disjointed REIPs occurring at a lower threshold for grape leaves, than for the various vegetation species of the Le Maire et al. (2004) study. In addition, the range of the gap was smaller, approximately 5nm (from 705-710nm), compared to a 25nm gap (Cho & Skidmore, 2006). With such a small data gap, it would be difficult to note without a high spectral sampling instrument. This corroborates the use of hyperspectral datasets and processing methods over generic broadband equivalents. This subtle gap feature affecting the REIP may have eluded detection until more recently. The other hyperspectral processing technique, continuum removal, also exhibited effects from the peak jump feature; the right area metric displayed increased variance at higher chlorophyll levels, which was stabilized if the

second peak group was removed. However, the reflectance indices did not capture the peak jump characteristics in their results.

Derivative methods that utilized the shape of the reflectance curve (e.g. DDI, AltREIP), rather than the location of the inflection point, were selected to aid in better quantifying the relationship between the reflectance slope and chlorophyll content. However, the waveband selection for the derivative models demonstrated that small variations, such as the difference between published and data-customized bands, had a significant impact in the improvement of the model. This was particularly the case in the DDI index, as three iterations of different band combinations were attempted, and all yielded different results. This suggests a sensitivity of the model that could be limiting if the importance of band selection is not conveyed. In addition to band selection, data range influenced the results. In that regard however, the DDI model had more success. The experimental dataset used to develop the DDI did not have very low chlorophyll levels**,** therefore, this dataset tested the limits of the model's robustness (Le Maire et al., 2004). This accounted for the higher percentage of negative values in the metric values compared to the original study. This large proportion of negative values can also be seen in the other two modified indices, R'mND and R'mSR.

By examining the difference between four locations (670nm, 700nm, 720nm, and 750nm), the DDI index provided a better sense of the changes and characteristics of the derivative than by simply examining the REIP peak alone. Besides the larger coefficient of determination, there is also a theoretical consideration: the DDI is a mathematical difference index, so it does not rely on the exact peak location, but rather several locations that contain different influencing effects within them. This broader reference

range is an advantage because the REIP 'peak' can sometimes be platykurtic (light coloured lines below), especially at higher chlorophyll levels (Figure 4.2). When the REIP spans 40nm, this peak value can start to look somewhat arbitrary. A final observation about the DDI index was that while Le Maire et al. (2004) only looked at the difference between the reflectance values at four locations, it would be of interest to see if the area under the derivative curve could provide any additional insights that were not captured by the difference alone. Only a few studies investigated the idea of an area derivative metric, though it was with little success (Elvidge & Chen, 1995; Murphy et al., 2005).

Cho and Skidmore (2006) had success by generating the AltREIP index from several

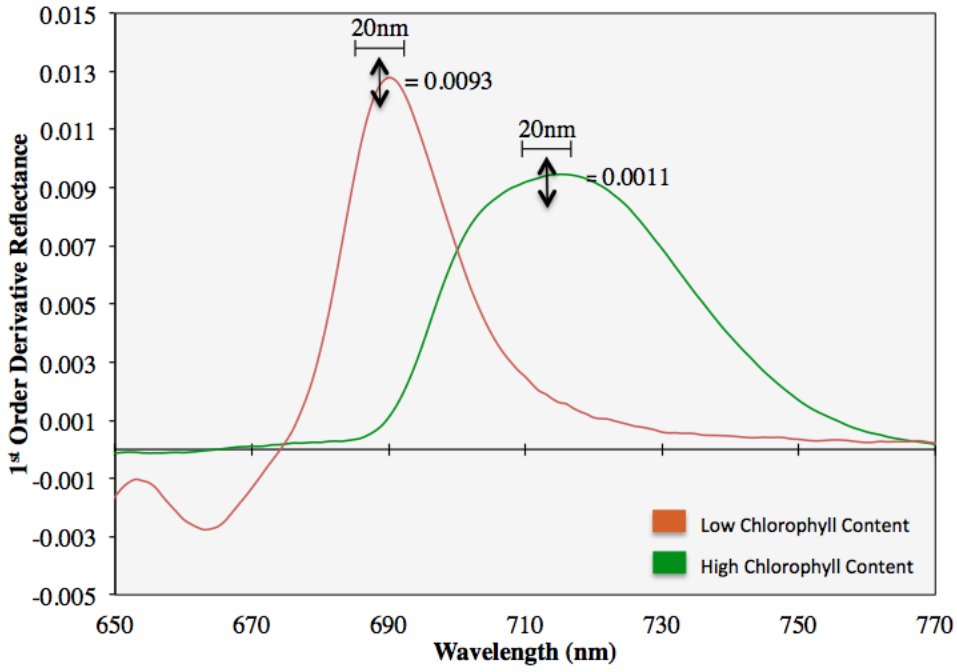

Figure 4.2: The platykurtic changes of the reflectance between the highest and lowest chlorophyll levels in the first order derivative data (green and brown, respectively). A distinct peak is noted at the lower (brown), but the vertical change greatly diminished over the same distance when the spectra shift to longer wavelengths. Values presented indicate the amount derivative reflectance change on either side of the REIP over a 20nm horizontal range.

spectral regions. However, this approach was unsuccessful at improving upon a single band model when applied to this study. While nitrogen estimations in their study yielded a strong linear relationship, results found it not appropriate for chlorophyll estimation. The AltREIP was hypothesized as a way to mitigate the effects of the double peak feature, and while successful at overcoming the discontinuity, the prediction model established in this study was not optimal for estimation. This was due to the inaccuracy of chlorophyll modeling at low levels. In addition, the residual patterns and dispersions indicated that the relationship was nonlinear and likely confounded by other unaccounted variables.

It was decided for this study that the derivative method would be used to determine the REIP; however it is noted that a few studies use multiple methods of determining REIP that do not necessarily rely on the first order derivative (Baranoski & Rokne, 2005; Cho & Skidmore, 2006). The focus of this section was the derivative methodology for chlorophyll estimation, rather than on the REIP feature, but the vast majority of derivative work focuses on the red edge. The overall success of the derivative was not very promising, but does not negate the red edge region of the electromagnetic spectrum as a poor chlorophyll estimate. Rather, it demonstrated the derivative was not an appropriate choice to utilize in this region, despite previous literature supporting the contrary (Baranoski & Rokne, 2005).

The results of this study suggest the REIP to be an inadequate estimator for chlorophyll. While this does not mean derivative data in general is not useful in the estimation of chlorophyll, as demonstrated by the success of the DDI, it does highlight that the REIP is not ideal in its isolated form when concerning fine spectral resolution

sensors. The theory that derivative data are better suited to identifying areas of interest rather than for generating data-based metrics was supported here.

4.1.1.3 Continuum Removal

The continuum removal (CR) approach captured the complete red absorption feature, and it follows that this feature is strongly related to the estimation of the quantity of chlorophyll present in the grape leaf. Although the performance metrics of the CR metrics were strong ($R^2 \geq 0.84$), not all of them were determined to be suitable models for chlorophyll estimation. The total area, left area, and right area are discussed, and the role of other foliar biochemical on these metrics is investigated.

The Total Area metric was the strongest estimator because it has a smaller standard error of the estimate than the left area counterpart. This was supported by its large F-ratio, reflecting the improvement of the model's explanatory power. In comparing the results to other imagery studies, the Total Area results were similar to Malenovský et al. (2006, 2013) with the $ANCB_{650\text{-}720}$ index, which was functionally the same as the Total Area metric. They cited a limitation of the index as an overestimation at lower chlorophyll content, however, this was not observed in this leaf level dataset.

The left side of the feature captures half of the red absorption feature and half of the green peak. The Left Area is limited as there is a saturation point that is a constraint for pigment estimation at higher concentrations (Mutanga & Skidmore, 2004). This saturation was responsible for the rejection of the maximum band depth (MBD) as a viable metric. Although having positional significance, MBD is ultimately impractical in chlorophyll prediction. Kokaly et al. (2007) outlined that the depth of the continuum removed feature is a function of the strength of the chlorophyll pigment absorptions and

their concentrations. However, the energy absorption saturation of chlorophyll pigments causes the absorption feature to plateau, and was responsible for the left area being characterized by a quadratic function.

While pigments exclusively influence the left area, the right area of the continuum captures the red absorption as well as the structural scattering and water influence associated and established in the near infrared region (Gitelson et al., 2006). Since the relationship is not pigment-exclusive, it is confounded and results in a heteroscedastic nature (confirmed through residual examination). While there are ways to correct for influences in the data, applying those corrections can inadvertently hinder the ability of the model's predictive power by forcing a data specific function, limiting the equation to descriptive rather than predictive. Such was the case with the right area concerning the heteroscedastic spread of data. To obtain the results in Table 4.1, a weighted least squares (WLS) was applied to the right area metric. WLS is designed specifically to reduce the

Table 4.1: Summary of the performance metrics results for the continuum removed index regression models.

Index Name	Coefficient of determination R^2	Standard Error of the Estimate RMSE	F-ratio
Total Area	0.93	± 1.96 µg/g	9535*
Left Area	0.90	± 2.44 µg/g	9533
Right Area	0.84	n/a (WLS)**	664

* indicates a calculated not reported statistic. In this case, F-stat was the ratio of the mean square model and residuals.
** indicates that an additional statistical approach was applied, rendering the metric incomparable to the other RMSEs

sum of squares by weighting heteroscedastic observations, and the reduced RMSE reflects this. However, because of this correction, it is not reasonable to compare to the other standard errors, nor to use towards extrapolations of other datasets.

One of the challenges working with this dataset was the inconsistency of the pigment types between leaf varieties. While it was demonstrated that all samples did have some amount of chlorophyll present, there were no quantitative methods explored to extract all pigment types. There was an assortment of leaf colour evident, resulting from differing pigment types and concentrations (Roberts et al., 2011). The reflectance spectrum will behave differently when there are non-chlorophyll pigments present in leaves (Qin, Rundquist, Gitelson, Tan, & Steele, 2011). Gitelson (2012) found that while leaves with no anthocyanin had a reflectance measurement closely related to chlorophyll at 550nm, leaves with anthocyanin and the same chlorophyll content had very different reflectance values, varying up to ten times. A study by Quinn et al. (2012), found an inflation of the CRRA due to a blue shift that was related to the presence of anthocyanin rather than chlorophyll content. Some of the more extreme samples in this dataset that exhibited clear high red pigment content could have potentially contributed to the increase of spread in the variance of the right area. The average leaf in the vineyard did not have high levels of anthocyanin, with striking red leaves being few and far between, as most varying in degrees of yellow to orange (Figure 4.3). The presence of some anthocyanin was supported by the correlogram results, as a local peak around 550nm suggested a weakened chlorophyll correlation; this could be explained by a strong anthocyanin absorption obstructing it (Thenkabail et al., 2011). While chlorophyll is the most influential and abundant pigment, the presence of other pigments is acknowledged.

Figure 4.3: Field photos depicting the varying degrees of leaf colours during the last sampling days. While the majority of leaves senesced to yellow-orange hues, occasional red leaves were observed for some varietals.

Results of this study aim to catch the robust trending behaviours of grape leaves during the senescence process. However, by extracting only chlorophyll *a* and *b* pigments, they are the only foliar biochemicals that can be definitively reported on. Methods of extracting both chlorophyll and anthocyanin together are discussed in section 5.3.

4.1.1.4 Validation

The validation of the reflectance had three main points of interest. First, the difference between the area and the fresh weight chlorophyll extraction validation results showed that fresh weight was a more appropriate metric when investigating chlorophyll in conjunction with non-destructive reflectance measurements. Secondly, it revealed that for the validation, the training dataset performed almost as well as the jackknifing, however, both methods had limitations. Neither method captured the higher chlorophyll concentration changes without a loss of accuracy when the CRTA reflectance index expressed chlorophyll in units of area. This was determined to be a limitation of the dataset rather than the process, and directly contributed to the chlorophyll per unit fresh weight being the superior unit of measurement. Lastly, the validation for the reflectance

index performed better than the SPAD validation, thus determining the spectrometer as the superior sensor to characterize chlorophyll content.

4.1.2 SPAD

Four main points were highlighted from the SPAD analysis. Chlorophyll expressed in different units was found to affect the ability of SPAD to estimate pigment content. The published SPAD equations applied to this dataset were unsuccessful at accurately estimating pigments, as they consistently overestimated chlorophyll content. SPAD equations were able to estimate chlorophyll content in grapevines to a threshold beyond which point SPAD modeling accountability decreased by a factor of 28%. Lastly, leaf physiological characteristics (moisture, light dispersion, and thickness) and sensor design limitations were identified as impediments of the SPAD's ability to characterize grapevine chlorophyll content.

Similar to Marenco et al. (2009) findings, a strong curvilinear relationship was established between chlorophyll meter and chlorophyll pigments. Comparing models, a higher coefficient of determination was obtained when chlorophyll was expressed in terms of leaf area than fresh weight, R^2 = 0.86 and 0.80. Despite the larger R^2 for chlorophyll by area, the angular deviation from the 1:1 ratio line (Figure 3.3a), comparing the estimated with the measured chlorophyll value, was more pronounced, with a larger underestimation of the higher chlorophyll values. Both Y-intercepts overestimated chlorophyll (196 μg/g and 3.44 μg/cm^2), however, these were small deviations compared to the standard deviation of the two datasets (366 μg/g and 5.97 μg/cm^2). The angular offset from the leaf area model remained after stratifying data to ensure full chlorophyll range characterization in both calibration and validation datasets, which undermined the

validity of using leaf area over fresh weight to estimate grapevine chlorophyll. These results were robust to experimental factors such as the data splitting for the calibration, as samples were the same for both area and fresh weight, and the same SPAD sample values were used for both chlorophyll measurements.

The overestimation in chlorophyll content with the Uddling et al. (2007) and Ceroivic et al. (2012) algorithms suggest that it is not appropriate to employ published SPAD equations for the estimation of chlorophyll for this grapevine dataset. Ceroivic et al. (2012) proposed a robust SPAD conversion equation comprised from multiple studies that would be insensitive to species variation. Markwell et al. (1995) noted that the SPAD transmittance measurement at 940nm compensate for differential scattering caused by varying leaf thicknesses and by association, species. However, the results revealed a bias in the data as a vertical shift between the measured and estimated chlorophyll values, resulting in a consistent overestimation of chlorophyll content within the grape leaves. The bias could be corrected for, if a constant could be determined to subtract from all estimated samples, but destructive pigment extraction would be required in order to quantify the bias. Extraction defeats the purpose of the fast, non-destructive chlorophyll meter, and generating an independent, data-specific model would be a better use of destructive measurements than a correction constant. The poor performance of these results indicate that the conclusions of Richardson et al. (2002) were correct in stating that SPAD predictive equations for chlorophyll should be species specific, and that a universal equation was not recommended, as structural attributes limit across-study applications. Upon performing a similar assessment of published equations to the one conducted in this study, Richardson et al. (2002) found comparable discrepancies; four

out of five published equations yielded unfavourable estimations compared to the measured chlorophyll of their dataset. The equations of Uddling et al. (2012) demonstrated that wheat, birch, and potato had varying results amongst themselves, and did not estimate grape leaf chlorophyll effectively.

The literature identified a threshold in chlorophyll content, beyond which the SPAD modeling decreases in accuracy, and a similar limit was observed in some of the results of the present chlorophyll meter analysis (Lin et al., 2015; Monje & Bugbee, 1992; Richardson et al., 2002; Steele et al., 2008b). Before this is expanded upon, the distinction must be made between model saturation due to the curvilinear relationships of independent and dependent variables, and model inaccuracy due to increased variance between the measured values and the model estimation along the regression line. A strong non-linear model can still be an accurate estimator of the dependent variable, with the caveat that often there is a threshold beyond which a change in one axis will yield no change in the other. Inconstant variance across the independent axis data range indicates an increase of point dispersion between the two variables, and results in non-uniform error distribution. A significant restraint on the use of the chlorophyll meter was that the estimates of leaf chlorophyll were less accurate as foliar pigment content increased. The jackknife results showed increased individual sample variance in comparison to the whole dataset, and it was seen to increase with chlorophyll content, meaning the SPAD estimation disproportionately increased with chlorophyll content, suggesting a threshold could be identified. The literature identified three different thresholds: 65 $\mu g/cm^2$ by Monje et al. (1992), Steele et al. (2008b) at 30 $\mu g/cm^2$, and Le Maire et al. (2004) at 45 $\mu g/cm^2$. It should be noted, however, that the destructive sampling of Steele et al. (2008b)

only included 20 samples for chlorophyll estimation, and that the CI Red Edge Index was used to determine the majority of the 288 chlorophyll values from their study. Given that the range of this dataset is approximately 0 – 35 μg/cm^2, according to the presented thresholds, no significant increase in variance should be noted. However, this was not the case. The residual plot from the regression of the known chlorophyll (by unit area) with the estimated chlorophyll (via the jackknifing technique) revealed that the increased dispersion can be seen as low as 25 μg/cm^2. Below this boundary the performance metrics were markedly better, $R^2 = 0.77$, RMSE = ±2.07 μg/cm^2. Above, the RMSE more than doubled to ±4.47 μg/cm^2 and the R^2 dropped to 0.49. Alternatively, the threshold could be determined as the point where samples started to consistently occur outside the prediction intervals (PI) for the regression model, in which case the threshold was set closer to 27 μg/cm^2. These results suggest that the dispersion was dependent on the range of the data; with a larger data range would come a larger dispersion and thus the PI would be wider, shifting the position where samples disperse beyond the PI's to a higher threshold. The dispersion was constant when considering foliar pigments expressed per unit fresh weight, and therefore heteroscedasticity was an issue specific to chlorophyll expressed in terms of unit area.

To contextualize the diversity of the SPAD modeling results, theoretical explanations were examined. There are three main factors that could account for error: overgeneralized models, sensor design limitations, and leaf physiological characteristics. These different characteristics can individually or collectively affect SPAD measurements, and are an amalgamation of explanations presented within the literature.

It was demonstrated that the chlorophyll meter required species-specific modeling. In addition to this, studies such as Taskos et al. (2014) alluded that temporal specific models also should be considered; they demonstrated that the chlorophyll meter is incapable of identifying the large chlorophyll drops at late stages of senescence, and continued to report inflated SPAD values. This rapid pigment decrease is characteristic of grapevines late in the phenological cycle (Strever, 2012). The omission of this change would result in models with an overestimation of chlorophyll content. However, instances of underestimation were observed in the jackknifing validation when chlorophyll was expressed by fresh weight (Figure 3.4). This underestimation was seen more so in the upper half of the chlorophyll range. The number of overestimated samples was not of concern, as they are within the 5% expected to fall two standard errors outside the 95% confidence intervals. Instead, what was of interest was that residuals favored under

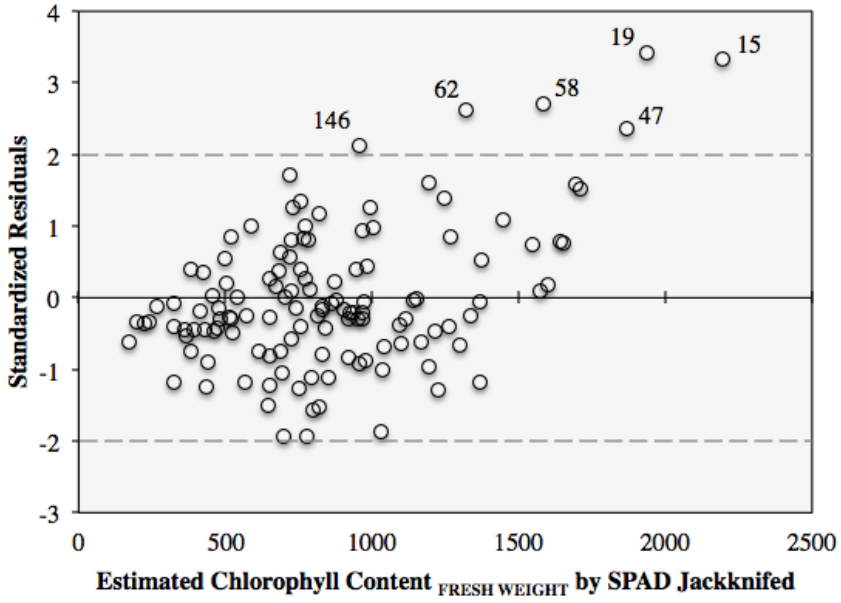

Figure 4.4: Scatterplot of standardized residuals from a regression of estimated (via SPAD modeling) and actual chlorophyll content (expressed in terms of fresh weight). Results show sample IDs that fall outside two standard deviations away from the zero-axis.

estimations rather than an even dispersion, as well as the strength of the deviation (the most prominent of which was almost four standard errors away) (Figure 4.4).

SPAD model improvement may be possible if results were segregated to correspond with major phenological developmental stages (Broge & Mortensen, 2002). The jackknifing of chlorophyll in units related to leaf surface area did not show overestimation, instead a fan shaped pattern occurred. Likewise, the residuals of the calibration and validation via data-splitting for both the area and fresh weight chlorophyll with SPAD did not indicate overestimation. Rather, an underestimation was more commonly occurring. Cerovic et al. (2012) placed emphasis on SPAD equations producing better results with mature agriculture crops, however, this was demonstrated not be the case. Regardless of whether over- or under-estimation was more prominent, as both were observed in these results, the literature suggests that a more temporally specific equation could provide a tighter and more accurate residual dispersion. This lends weight to the hypothesis that senescence plays a role in the SPAD modeling.

Depending on the chlorophyll range, it could be suggested that some studies were attempting to push the capabilities of the instrument beyond its intended purpose, as they are insinuating relationships beyond the sensor range (Figures 4.5). There are evidently vegetation species, particularly within agriculture crops, that surpass the chlorophyll content with which the SPAD was intended to model. Several studies reported that the SPAD does not report accurate readings past a certain chlorophyll content (~30 $\mu g/cm^2$); however more emphasis is placed on the chlorophyll range than the SPAD value range (Le Maire et al., 2004; Monje & Bugbee, 1992; Steele et al., 2008b). The SPAD user manual states that it is not able to take accurate readings past 50 SPAD units (Spectrum

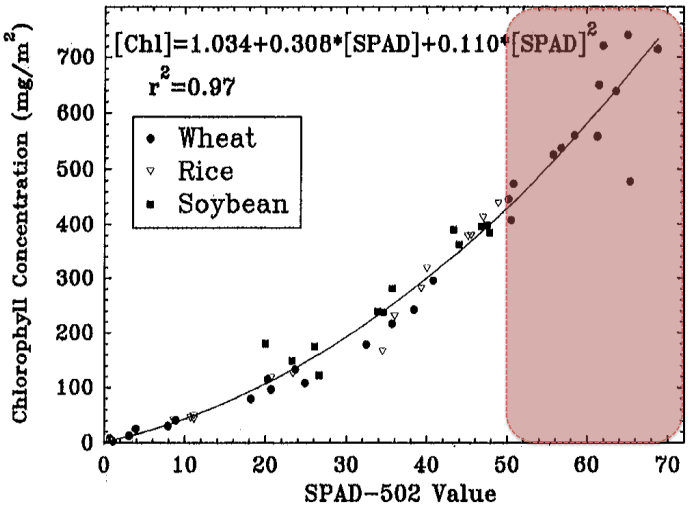

Figure 4.5: Graph adapted from Monje and Bugbee (1992). Author's identified threshold of 600mg/m^2 beyond which SPAD readings were overestimated. Red highlighted polygon was added at SPAD value = 50 to indicated the meters accuracy limit. This demonstrates potential sensor unsuitability for some vegetative crop types.

Technologies Inc., 2009). The range limitation was not evident with this dataset pertaining to senescing grape leaves; there was only one sample where the averaged SPAD readings reported a value higher than 50. Results from this study qualify the SPAD as an appropriate sensor, as the chlorophyll meter's design was not a limiting factor for the analysis in this senescence study in regards to the chlorophyll range boundary. However, different sensor features, such as the transmittance bands selected to determine the SPAD value, contributed to the effectiveness of the SPAD modeling, and are discussed next in terms of leaf characteristics.

Heteroscedasticity is often encountered in regression residuals as the result of an additional variable influencing the relationship between two variables that has gone unaccounted for (Hayes & Cai, 2007). 940nm is the NIR band used in the chlorophyll meter calculation for the generation of the unitless SPAD value. While it was intended to capture the function of cellular structure as a contrast to pigment influences, this band

falls within a water absorption feature. Kokaly et al. (2007) identifies 980nm as being the centre of the absorption feature, while Pu et al. (2003) indicated 975nm as the centre, with an absorption range of 920-1120nm for oak leaves. Sims and Gamon (2003) stipulated this was a broad absorption features with a centre at 970nm. Regardless of the specific location of maxima (which will vary slightly with vegetation type), the consensus stands that 940nm is on the shoulder of this water feature, and therefore is sensitive to changes in cellular water content (Figure 4.6).

Rather than compensating for scattering, the 940nm band adds an additional influencing factor into the determination of the SPAD values. Lin et al. (2015) demonstrated that water stress could lead to overestimation in chlorophyll content if a model was developed with sufficient moisture and then applied to water deprived plants, and Marenco et al. (2009) and Martinez et al. (2004) shared this finding. This is due to a

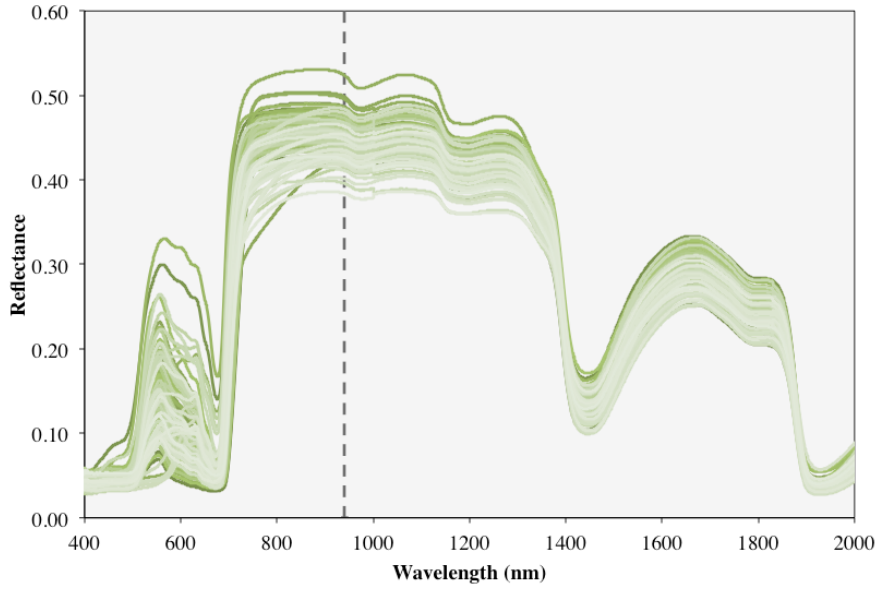

Figure 4.6: All samples from this study identifying water absorption feature at ~960nm to demonstrate the location of 940nm on the shoulder of the water feature.

reduction of the light path length leading to reduced absorbance with the potential for increased transmittance and scattering.

Regarding the other bands utilized in the internal SPAD calculation, Cerovic et al. (2012) suggested the use of 710nm instead of the 680nm could improve the modeling ability for transmittance measurements. This region is not affected by other pigments, and not at risk from other foliar pigment influences overshadowing the chlorophyll meter readings. While using band 710nm provided a more linearized relationship, as it is a band located closer to the maximally correlated waveband position of 702nm, the use of this band still exhibited similar limitations (i.e., heteroscedasticity) that the 680nm band displayed (Cerovic et al., 2012).

When light interacts with a leaf, pigment abundance, pigment distribution, internal light scattering, and leaf surface reflectivity determine the transmittance of light (Monje & Bugbee, 1992; Vogelmann, 1993). Under these four categories, further influences can be characterized; for example, light scattering is subject to internal cell arrangement and the refraction between the interfaces of cell walls and air (Zhang, Chen, & Thomas, 2007). The pigment distribution determines the extent of the sieve effect and the detour effect. The sieve effect is one of the more commonly addressed leaf characteristic explanations within the literature to account for the increased variance of SPAD measurements in relation to larger chlorophyll content. It centres on the concept of pigment dispersion within the leaf thickness, and the degree of distribution uniformity. Chlorophyll pigments are housed in chloroplasts, and increased abundance of these housing units cause more pronounced concentration differences throughout the leaf if chloroplasts are not uniformly distributed (Marenco et al., 2009; Markwell et al., 1995;

Monje & Bugbee, 1992; Ustin et al., 2009). The sieve effect describes the instance of light transmitting through the leaf microenvironment without encountering a highly concentrated chloroplast, resulting in the characterization of a different chlorophyll content than the plant may truly have (Markwell et. al., 1995). Counterpart to the sieve effect, the detour effect describes the increase in absorption from scattering and the extension of the path length, resulting in lower transmittance of light through the leaf compared to a uniformly distributed microenvironment (Vogelmann, 1993). However, the resulting outcome of the sieve effect is not universally agreed upon. Richardson et al. (2002) stated simply that negative effects would result and performance would be poorer at higher chlorophyll contents. Marenco et al. (2009) concluded non-uniform pigment distributions lead to lower SPAD values, while Monje and Bugbee (1992) determined the sieve effect would yield overestimation of SPAD values at higher chlorophyll contents. Markwell et al. (1995) placed more significance on the detour effect rather than the sieve effect that would result in the overestimation of chlorophyll from the SPAD meter.

Marenco et al. (2009) presented another possible leaf characteristic that influenced SPAD values; they determined that SPAD values increased with fresh leaf thickness, which in turn would reduce transmittance. The effect on the SPAD readings would be recorded as more absorbance from pigments. Adding to this, Koundouras, Tsialtas, Zioziou, and Nikolaou (2008) noted that grape leaf thickness increased with age. Following this logic, SPAD values should increase with grapevine phenological development, and consequently an overestimation of SPAD values compared to the measured chlorophyll content would be the expected observation. Adding further, the ultrastructure and dispersion of chloroplasts within grape leaves change substantially

during senescence, which could contribute to any over- or under- estimation, or general discrepancies that may be observed (Bertamini & Nedunchezhian, 2003). Without a series of more testing and sampling with controls and known variables in place, the results of this study cannot be attributed with certainty to any one of the above effects. All that can be definitively said is that the light-scattering properties and cellular interactions within the leaf interior appear to constrain the ability of the SPAD chlorophyll meter to estimate in vivo chlorophyll concentration (Monje & Bugbee, 1992).

All of the above factors influenced and constrained the chlorophyll meter's ability to estimate chlorophyll content accurately. This was particularly the case at the higher chlorophyll levels. Despite the simplicity of SPAD operation and collection, and the seemingly straightforward unitless value it generates, there is still much to be investigated regarding the SPAD in relation to non-destructive chlorophyll estimation. There are discrepancies and contradictions within the literature, with emphasis placed on a variety of different determinants to account for the inconsistencies. There are continued efforts to investigate the relationship between chlorophyll and SPAD measurements, with more recent studies diving into discerning the individual chl *a,* chl *b,* and total chlorophyll effects on SPAD values (Ibrahim & Jaafar, 2013). Studies like that of Nauš, Prokopová, Řebíček, and Špundová (2010) are investigating the effect of the time of day and the daily migration of chloroplasts in relation to changing SPAD readings. The results even within this study displayed a range of effectiveness of the SPAD depending on the comparing variable and the processing of data. All these factors combine to demonstrate that more research is needed for the SPAD as a chlorophyll meter as well as

a tool to remote sensors attempting to collect field measurements to accompany their airborne data collections.

4.2 General Discussion

Influencing Factors of the Datasets

In evaluating two non-destructive sensors, a number of different data processing methods for remotely sensed data were reviewed. Reflectance and transmittance metrics for grapevine remote chlorophyll estimation were assessed. Individual models varied depending on the sensor, the data manipulation, and the spectral region used in analysis. Some of the model success or failure were attributed to these three factors, and were discussed in section 4.1. A general trend emphasized in the results was heteroscedastic residuals; they occurred in the jackknifed validation SPAD results (chlorophyll expressed in terms of area), as well as both published SPAD indices (chlorophyll expressed in terms of area). Inconstant variance also was observed for the data-split validation of the CRTA reflectance metric (estimated versus measured chlorophyll expressed in terms of area). While the majority of inconstant variance occurred with the use of area measurements, the CCRA also had heteroscedastic residuals (chlorophyll as fresh weight). More fundamentally, there was also inconstant variance between the original chlorophyll measurements when expressed in different terms.

This discussion investigated possible contributing factors to heteroscedasticity, as there was likely more than one major influence. In addition, one of the goals of this discussion was to determine if abnormalities in the results were caused by plant physiological factors, the result of a statistical method, or the sampling strategy. Different topics were investigated including (i) leaf moisture, (ii) pigment sampling design and the

impact of leaf density on sampling, (iii) the role of senescence and data distributions, and (iv) data transformations. Regarding the leaf physical characteristics, no one variable accounted for the majority of residuals dispersion, but because they were not controlled for in the study, the exact extent of their influence currently is indeterminable with the given datasets. More thorough testing to conclusively verify would be required to ensure results were not spurious. The following were theoretical considerations.

The first possible influence was leaf moisture levels. There is a well established link between leaf water stress and chlorophyll content, and was a primary concern for many agriculture crops studies (Chutia & Borah, 2012; Rodríguez-pérez et al., 2007). It was hypothesized that leaf moisture was an additional variable in the relationship between non-destructive measurements and pigment content. As previously noted in section 4.1.2, it was suggested that the SPAD measurements were likely susceptible to changes in moisture, as the second transmittance measurement used in generating SPAD values was located on the shoulder of a water absorption feature. However, water content was determined not to have played a significant role in effecting the pigment-SPAD measurements. While the chlorophyll measurements increased in variance with higher SPAD values, the water content remained comparatively constant (Figure 4.8a). This was seen also when moisture was compared directly to chlorophyll pigments expressed in terms of fresh surface area (Figure 4.7).

Moisture variance appeared smaller at higher chlorophyll contents. To further investigate this, the residuals of the SPAD-pigment model were graphed with the moisture measurements to assess if the error variance was correlated to moisture. However, no such relationship was found, as the dispersion of residuals did not present a

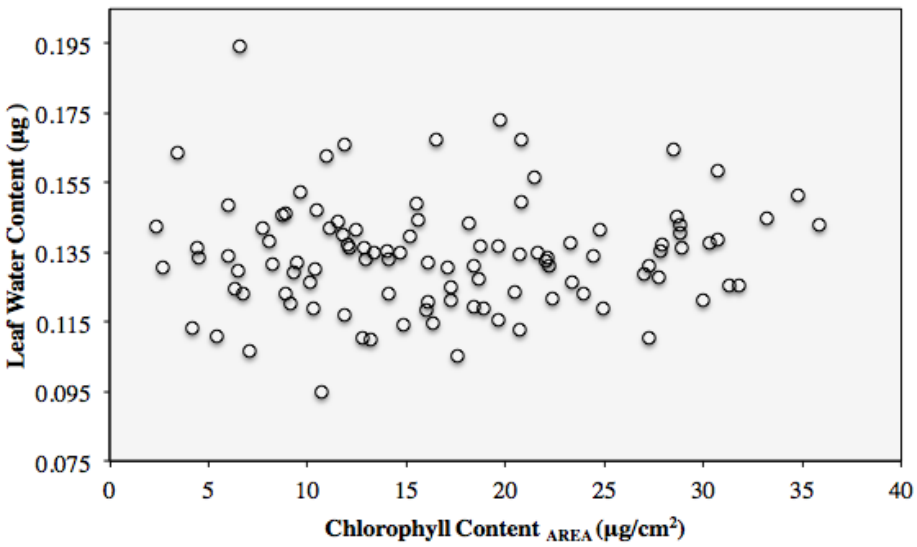

Figure 4.7: Comparing water content of grapevine leaf samples with chlorophyll content of the samples expressed in terms of surface area. Water appears constant, with a smaller variance at higher levels. This does not support the theory of water significantly influencing chlorophyll relationships.

pattern to indicate an influential trend (Figure 4.8b). When modeled, the unstandardized absolute values of the residuals obtained a weak positive relationship to moisture, R^2 = 0.03. Based on these results, only approximately 3% of the variation in the variance of the chlorophyll-SPAD model was attributed to moisture variation. However, because variables were not controlled for, this could be a spurious relationship.

The second factor investigated for influencing the results of this study was the method of determining chlorophyll pigment content. It was assessed for potential sampling and/or design bias (Pannucci & Wilkins, 2010). Results showed that metric performance changed depending on whether chlorophyll was expressed in terms of fresh leaf weight or fresh leaf area. Specifically, results of both the jackknifing for the SPAD and reflectance metrics differed in this regard. The fresh weight was kept constant at 100mg, but the area varied for each sample, so the number of leaf cores used in extraction

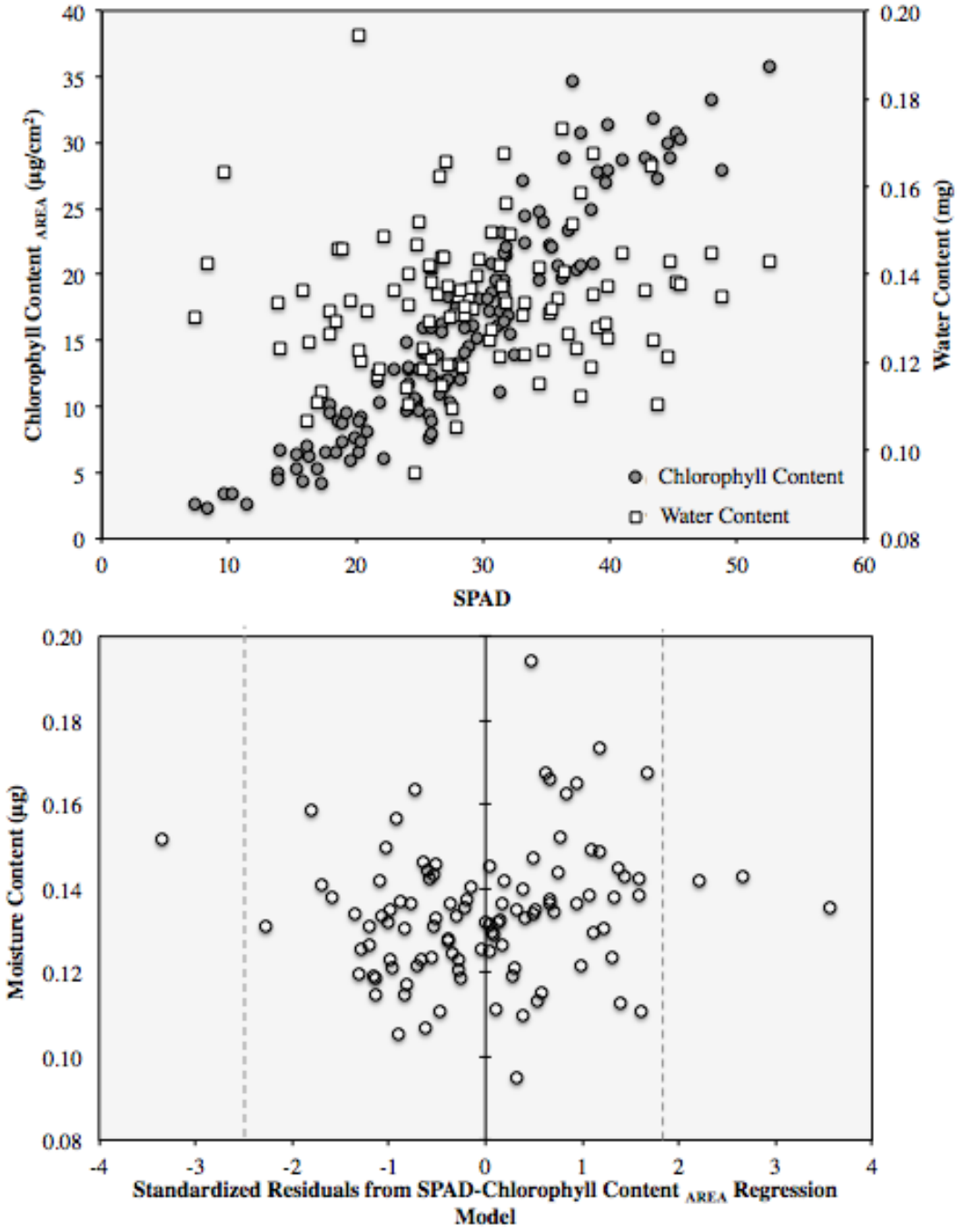

Figure 4.8: (a) scatterplot depicting the difference between chlorophyll$_{area}$ and leaf water content with regards to varying SPAD measurements. Chlorophyll$_{area}$ (dark circles) increased with SPAD values, with a noticeable increase in variance at the higher values. Comparatively, leaf moisture content remained fairly constant with regards to SPAD values, suggesting water is not the variable responsible for the pigment variance increase. (b) A uniform random spread of the resulting unstandardized residuals from SPAD-pigment modeling (along the X-axis) showed no pattern when graphed against leaf moisture (Y-axis), which further suggested there was no significant relationship between moisture and the unaccounted heteroscedasticity.

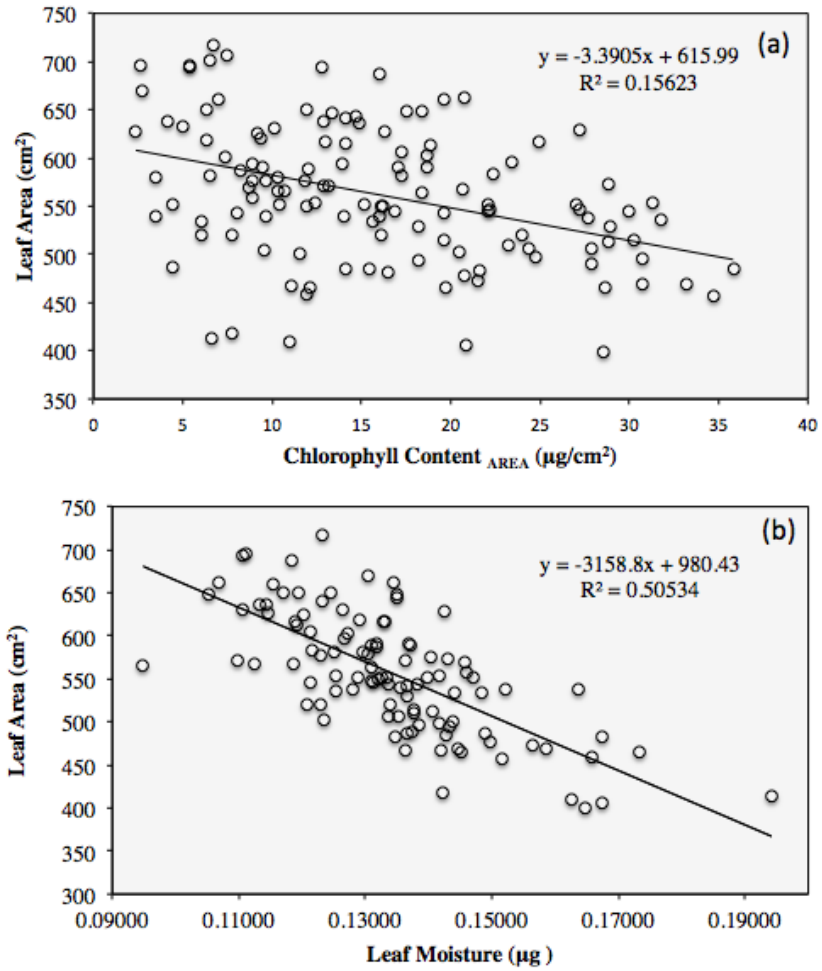

Figure 4.9: Grape leaf area measurements graphed against (a) chlorophyll in terms of area and (b) leaf moisture. Results indicated a negative correlation between leaf area and moisture, and is supported by the theoretical explanation that if weight remained constant for sampling (0.0100g), less moisture would be required in the samples that had larger total areas to maintain a fixed weight.

also varied. Figure 2.7 confirmed that the variance between both chlorophyll contents was not constant, as the discrepancies between the higher chlorophyll contents were larger compared to the lower chlorophyll levels. The increased variance was determined not to correspond to the larger leaf areas; in fact, the opposite was observed. The samples with the highest number of cores (i.e. largest surface areas) had lower chlorophyll levels (Figure 4.9a). Therefore it was established that the uncertainty accumulating with each leaf disk area measurement was not compounding to significantly vary the dispersion

between chlorophylls (more disks did not equate to more error). There were two possible ways to interpret the results of the difference between area and fresh weight chlorophyll measurements.

The first interpretation centred on leaf volume and densities. It was based on the theoretical principle that a fixed leaf weight with a larger area would result in more volume (assuming the leaf thickness was kept constant). This means that leaf density would have to drop to compensate for the volume increase to maintain the same weight. A lower density translates to more air and less water, thus lowering moisture content within the leaf.

This observation between leaf area and leaf moisture was seen in the results (Figure 4.9b). There was generally lower moisture content in samples that had a larger total leaf area for a fixed weight. Although moisture did not directly affect the chlorophyll, the moisture influence on leaf area may have contributed to the difference in the area and fresh weight measurements. However, this could not be confirmed as the thickness of the grape leaves were known to vary depending on the varietal of the vine, which would directly vary leaf density.

The second interpretation of the difference between the two chlorophyll measurements was that they indicated sun and shade leaf samples in the dataset. This supported the first interpretation, as it pertained to leaf structure and internal distribution of leaf layers, and thus leaf density. Shade leaves are thinner and have a larger mesophyll air space (Gausman, 1984). They have more chlorophyll on a mass basis, but less on an area basis (Strever 2012, Table 5). This was more evident for older leaves than younger ones. Shade leaves compensate by maximizing distributions more over area rather than throughout the

leaf to maximize exposure on the surface. These changes occurred through the palisade cell arrangement (Cui et al., 1991). Relating these findings to the current study, a combination of shade and sun leaves may have been sampled for this study, despite all samples having been taken along the same spurs at the top of the grapevine canopy.

An additional factor that potentially influenced the study's overall results was senescence. The process of modeling chlorophyll pigments to non-destructive measurements was confounded by the phenological stage of the leaves when the sampling occurred. This transitory state was more complex to characterize than a stable biological growth stage. In addition to differences in leaf thickness due to leaf placement in the canopy (sun or shade), leaf thickness also changed with age, and became thicker later in development (Strever, 2012). This was a result of increasing intercellular space. While leaves under stress with a decrease in chlorophyll may not lead to significant absorbance changes, older leaves do, as higher reflectance will occur in the visible and initially NIR, though the latter does decrease when the mesophyll collapses (Strever, 2012).

The changes from leaf phenological progression, especially the differences in spectral response in relation to pigment degeneration, effected the data distributions. Skewness reoccurred throughout the dataset frequency distributions causing non-normality, and was investigated thoroughly. It was rejected as measuring error, as the chlorophyll pigment values were within accepted ranges: 300-1970 μg/g for this senescent grapevine data, and a range of 700-2500 μg/g fresh mass was expected for healthy vines (Strever, 2012). Since collection occurred during senescence, it was probable that this accounted for the skewness; a representative range of the higher chlorophyll content was not captured

accurately, and because of the low frequency, appeared more as outliers than within the normal distribution range of the senesced dataset. The centre of the normal distribution shifted to lower chlorophyll measurements during this later part of the season.

The last examined potential influencing factor was the square root transformations applied to the regression variables prior to running the analysis. They were determined not to significantly contribute to the heteroscedasticity seen in the results. Initial concern was that the conversion back to original chlorophyll content units may have caused the non-constant variance, as the larger values increased more than the smaller ones, presented as a fan shape spread. However, this was disproven, as the heteroscedasticity was present whether the variables were transformed or not (results not presented, but tests conducted). This finding was supported by the fact that heteroscedasticity was present at the most basic level of chlorophyll measurements (between area and fresh weight). Further testing would be required to determine the exact extent of the transformation impact. This also confirmed that the coefficient of determination was influenced by data range. Assuming the standard error of the model estimate remained the same, increasing the data range improved the R^2 (Fadock, 2011). Therefore, even higher performance metrics could have been achieved if the non-square rooted values were used; however, the square root was used in favour of a more normal distribution.

Few remote sensing studies reported and discussed heteroscedasticity in their results (Murphy et al., 2005), yet it can be seen in other studies. For example, Gitelson and Merzlyak (1997) appeared to have heteroscedasticity in their validation results (Figure 4.10), which could have been confirmed by a Goldfeld-Quandt test. Unequal variance also was seen when modeling SPAD measurements with destructive chlorophyll

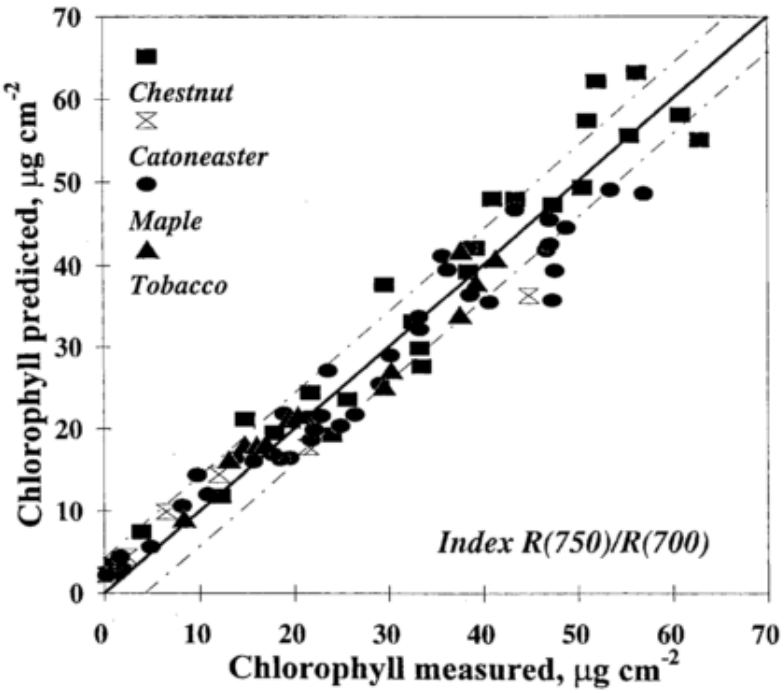

Figure 4.10: From the study of Gitelson and Merzlyak (1997), the validation of destructively measured chlorophyll of four vegetation types (n=96) against estimated chlorophyll via the R_{750}/R_{700} index. The solid line presents the 1:1 ratio, and the dotted lines are the RMSE. Although not reported, results indicate the variance may have been heteroscedastic, as indicated by the increased spread of catoneaster and chestnut samples at higher chlorophyll contents.

measurements of multiple vegetation types (Cerovic et al., 2012, second graph in Figure3).

There were many possible factors influencing the underlying relationship between pigment content and non-destructive measurements. However, the objective of this study was to determine how well that relationship could be modeled, regardless of the intervening variables. The data facilitated answering this, but with the caveat that they were not fully characterized due to other contributing complex factors. Additional analysis controlling for more variables would have to be conducted to statistically quantify causal factors. Nonetheless, strong models were determined within the abilities of the datasets.

Chapter 5 Conclusion

5.1 Summary and conclusions

This study performed a series of chlorophyll estimation models through an iterative regression process, followed by two methods of model validation. Reflectance indices, derivative indices, and continuum removed metrics were systematically explored and compared. Eighteen reflectance-based attributes and one transmittance-based attribute were used to generate predictive metrics, 15 of which were reported in this study. Performance metrics indicated strong models with high explanatory power, and provided contextualized results when compared with results from similar literature. The continuum removed depth normalized total area metric was presented as the optimal non-destructive attribute for accurate chlorophyll estimation for field campaigns ($R^2 = 0.93$). That said, the chlorophyll meter presented compelling results ($R^2 \geq 0.78$), and both non-destructive sensors were determined to be appropriate for field validation campaigns. The analysis discussion presented potential influential factors on the underlying relationship between non-destructive measurements and vegetation pigments. It concluded that, although the models were sensitive to pigment changes, there were confounding variables contributing to the relationship complexity, as indicated by the model residuals.

The quantitative study was designed to generate robust predictive algorithms, and the results indicated that a balance is required between robustness and specificity; models performance could be improved if refined by factors such as time or species, at the cost of model versatility and applicability. Ultimately, if a versatile model is to be developed, further studies would be needed to focus on additional leaf components, and environmental variables considered and accounted for (Kokaly, 2001). Accordingly, the

models generated in this study should be considered explanatory (rather than predictive) in an effort to better characterize the spectroscopic estimation of grapevine biochemical content, and that the validation strategies added to their explanatory power. This was also reinforced in other studies, as it was demonstrated that caution must be exercised in any predictive estimates beyond the specific data range used in the development of the model (McKillup, 2012). Nonetheless, consensus within the literature finds that non-destructive field measurements allow for quick estimation of pigment content and thus the physiological performance of leaves, and provided temporally-rich data collections for better future agriculture monitoring (Gitelson, 2012; Richardson et al., 2002; Ruiz-Espinoza et al., 2010).

5.2 Research contributions

This work contributes to the remote sensing research community, as the laboratory results add to the foundational knowledge necessary to establish valid remote sensing studies for vineyards. There is merit in approaches utilizing both chlorophyll meters and spectrometers, and their effectiveness has been empirically demonstrated. This study presented a novel opportunity to study the influence of biochemicals on the spectral response of grapevines as it pertains to the extreme ranges of the grape growing belt (the term given to the global region between 30° and 50° north and south of the equator, May & Sharpe, 1997). Given that there are over 35 vineyards on Vancouver Island, the detailed monitoring of these crops is essential, as the decisions made in the vineyard will determine the success of the wine (Hynes, 2011). Due to the geographic location of these northern coastal vineyards, it is likely that the region is more susceptible to changes in climate and will be affected before their interior or southern counterparts. The Venturi-

Schulze vineyard is a working model for sustainable ecological practises, and could be utilized as an example for policy making. Having successfully modeled chlorophyll content at a leaf-level scale, this information could contribute to future developments of spatial and temporal monitoring through a large-scale remote sensing approach.

The study provided information through the research objectives: the quantitative analysis succeeded in modeling foliar pigments to generate non-destructive estimates by both reflectance and transmittance field measurements. The principle objective was to develop a comprehensive characterization of the spectral responses during the dynamic senescing process undergone by grape leaves as a result of chlorophyll content degradation. In addition, a deeper understanding of leaf biochemical influences, and their effects, was cultivated and examined. The second objective involved a comprehensive analysis of established spectral processing methods, as well as an investigation of different statistical processing approaches. The latter revealed that simple ordinary least squares regression was capable of capturing strong causal relationships between pigments and hyperspectral metrics. The third objective established the ASD spectrometer superior to the hand-held chlorophyll meter, although both performed satisfactorily for grapevine chlorophyll estimation. However, a recommendation for future studies would be to let research objectives determine sensor suitability; the spectrometer provides much more additional information on other biochemicals and non-pigment plant constituents, while the SPAD is extremely portable and simpler to operate than the ASD spectrometer system.

The last objective of the study revealed and expanded upon several findings that contributed to current efforts in foliar biochemical estimation. The first was support for

phenological-specific models. The results, specifically the range and skewness of the datasets and subsequent effects on modeling, demonstrated a wide variation of responses that suggest it may be possible to generate specific health and senescent models separately. The late-stage variation was more difficult to characterize due to the rapid chlorophyll degradation in preparation for vine dormancy. However this phase is of particular interest for viticulturists, as the remote detection would provide vegetation information during the critical fruit ripening and harvesting period (Bramley & Hamilton, 2004; Serrano, González-Flor, & Gorchs, 2012). A second finding pertained to the chlorophyll meter, as previous research on chlorophyll model estimation using the SPAD was limited because pigment was expressed in terms of area. In this study, it was demonstrated the use of fresh weight provided more stability to the models, despite not achieving the highest coefficients of determination. Furthermore, this research raised concerns on identifying thresholds for chlorophyll estimation limits based on SPAD measurements: it was found that some thresholds were attributed to sensor capabilities rather than operational restrictions. The last contribution was to the developing body of work that questions the appropriateness of the REIP position as a chlorophyll estimator (Curran et al., 2001; Huang et al., 2004).

5.3 Research considerations and future directions

The limitations of this study have been recognized and recommendations have been compiled to aid any future studies examining non-destructive grapevine pigment estimation. In considering the model residual error variance results, there is potential for future studies of the same scope to include other plant and environmental factors to more thoroughly assess spectral contributors, as well as additional data extraction and

processing methods not included in this study. A logical extension of this study would be to increase the scope to include true remotely sensed non-destructive ground measurements (versus the current datasets of chlorophyll meter and contact probe measurements), which would require quantifying upscaling issues. Any future studies that establish stable relationships between vegetation spectral responses and fruit quality will be of interest. However, a current impediment to an airborne extension is the obstruction of fruit from sensors for direct measurement. Lastly, to facilitate remote sensing grapevine studies, unmanned aerial vehicles (UAV) technologies are recommended for field methods as they compliment the high spatial and spectral resolution requirements for detailed vegetation characterization. In addition, they are cost effective for high frequency operations to provide temporally rich datasets (Matese et al., 2015; Xiang & Tian, 2011). These topics are individually discussed in more detail below.

Future field-level studies that include more complete analyses of plant biochemical and non-pigment constituents may overcome the problem of model residual error through a more comprehensive multivariate approach for better spectral characterization, although would not be without its own set of challenges and limitations (Curran, 1989). It would, however, contribute to the understanding of the factors influencing the pigment-reflectance relationship. Nitrogen, lignin, and cellulose concentrations (Kokaly, 2001), as well as carotenoid pigments (Blackburn, 1998; Datt, 1998) are all plant components previously investigated, and could have provided additional information for this study. Specifically, given the data for this study were collected over the senescing period, and some grape varietals exhibited the typical red leaf hue to indicate the presence of anthocyanin, this particular pigment would be

beneficial for sampling studies during this phenological stage. This accessory pigment absorbs around the green peak (540-550nm), and thus could potentially affect the CRLA metric, as well as any other indices which utilize this spectral region (Gitelson, 2012).

DMSO was used in this study with the extraction equations presented by Wellburn (1994) to determine total chlorophyll. To integrate anthocyanin analysis into senescent studies, both chlorophylls and anthocyanin pigments can be isolated from extraction solvents using a pheophytinization process (Allain, 2007; Lashbrooke et al., 2010; Steele, 2007). This multi-step approach would first record the absorption of chlorophyll pigments, which are then converted through an acidification process to pheophytins, the result of magnesium break down (Lin et al., 2015; Strever, 2003). While anthocyanins are colourless in the neutral solution, the acidification develops the absorption feature which also then can be recorded (Steele, 2007). The complete extraction of additional plant biochemical components would facilitate more in-depth spectral studies, and possible errors in chlorophyll absorption could be quantified.

A final suggested direction for leaf-level research would be the expansion of models investigating the effects of individual species on spectral responses. While discrimination between species is of great interest within the remote sensing community (Cho, Sobhan, Skidmore, & de Leeuw, 2008; Cochrane, 2000; Schmidt & Skidmore, 2010), cultivar discernment within a given species has not been explored as thoroughly (Marenco et al., 2009). Preliminary results indicated that a degree of spectral separability was possible between the nine grape varietals within the study site (Parton, Quinn, Visintini, & Niemann, 2012). Cultivar discernment could also extend to chlorophyll

meter modeling as well, although it is anticipated to be more difficult due to the lack of detailed characterization captured in the SPAD readings.

The recommendations above are pertinent to leaf-level studies. However, in order to better aid the wine industry, grapevine studies will have to expand to produce data products at vineyard-wide scales that convey meaningful information to vineyard managers. This will allow spatial variability tracking and health assessment through pigment concentration determination (Borengasser et al., 2008). In order to develop estimation algorithms for true remote sensing data, the differences between laboratory and remote measurements have to be examined and quantified (Kokaly & Clark, 1999). Metrics generated from relationships at leaf-level may not be appropriate at an airborne scale, as a multitude of additional factors, such as canopy structure and soil background, compound to create complex responses (Gao, 2006). As well, the methodologies associated with spectral image processing should be considered because of increased dynamics associated with upscaling (Zhang, 2011).

There has been limited literature concerning grapevines and the remote determination and monitoring of crop quality. While a few studies linked vegetation physical characteristics (canopy area and density) to fruit quality parameters (ripening rate and berry size), grapevines are a unique crop in that they are often pruned multiple times throughout a season. This would affect the natural fluctuations and variations in canopy features (Hall et al., 2002; Johnson, Roczen, Youkhana, Nemani, & Bosch, 2003). In general, there is a limitation to remote estimation of grape quality attributes, as they are not directly measurable from airborne platforms, and their association with canopy characteristics are complex and only indirectly determinable (Taskos et al., 2014).

The direction of future vineyard studies will be driven by the development of remote sensing technologies to support the research; current UAV advancements are very promising in aiding small-scale precision agriculture work. Data collection by UAVs is less expensive than by rotary or fixed-wing aircraft, making it amenable to repeated measurements for robust temporal studies. The smaller size of drones allows the collection of high spatial and spectral information at low altitudes, decreasing the amount post-processing using spectral unmixing (Matese et al., 2015). While some UAVs are single sensor platforms that can derive multiple datasets (e.g. RGB imagery mosaics, and structure from motion, SfM, terrain models), there is also a generation of UAVs being developing that is capable of multi-sensor payloads (Mathews & Jensen, 2013; Saari et al., 2011). They would produce datasets akin to the current airborne collections. While this study investigated only one branch of remote sensing (the collection of optical reflectance in the visible and NIR region with passive sensors), active sensors can collect spatial terrain and vegetation information (Arnó et al., 2013; Rosellpolo et al., 2009; Viau, Jang, Payan, & Devost, 2005). Light detection and ranging (lidar) provides ancillary spatial information that can be used to further characterize canopy features related to spectral returns. The future directions for precision vineyard management, especially on Vancouver Island, are promising if remote sensing technologies are integrated with current practises. With a scientific aerial imaging industry targeted to vine management already established in the Californian wine regions (Staid & Staid, 2009), evolving studies can continue to contribute to the knowledge of grapevine spectral responses for better, more consistent yields in addition to improved sustainable agriculture practises.

Reference List

Allain, K. (2007). *Extraction of chlorophyll using dimethylsulfoxide and acetone*. Unpublished manuscript, Worcester Polytechnic Institute, Massachusetts. Retrieved from https://www.wpi.edu/Pubs/E-project/Available/E-project-042607-130620/unrestricted/KURTIS.pdf

Almeida, T. I. R., & Filho, D. S. (2004). Principal component analysis applied to feature-oriented band ratios of hyperspectral data: A tool for vegetation studies. *International Journal of Remote Sensing*, *25*(22), 5005–5023. doi:10.1080/01431160412331270812

Analytical Spectral Devices Inc. (2002). *FieldSpec Pro user's guide: ASD part #600000 Rev. C*. Boulder, CO. Retrieved from http://support.asdi.com/Document/FileGet.aspx?f=600000.PDF

Arnó, J., Escolà, A., Vallès, J. M., Llorens, J., Sanz, R., Masip, J., … Rosell-Polo, J. R. (2013). Leaf area index estimation in vineyards using a ground-based lidar scanner. *Precision Agriculture*, *14*(3), 290–306. doi:10.1007/s11119-012-9295-0

Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidases in Beta vulgaris. *Plant Physiology*, *24*, 1–15.

Atwell, B., Kriedemann, P., & Turnbull, C. (Eds.). (1999). Light use and leaf gas exchange. In *Plants in action; Adaption in nature, performance in cultivation*. Melbourne, Australia: Macmillan Education Australia Pty Ltd.

Azia, F., & Stewart, K. A. (2001). Relationships between extractable chlorophyll and SPAD values in Muskmelon leaves. *Journal of Plant Nutrition*, *24*(6), 961–966. doi:10.1081/PLN-100103784

Babyak, M. A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine*, *66*, 411–421. doi:10.1097/00006842-200405000-00021

Baldy, R., DeBenedictis, J., Johnson, L., Weber, E., Osborn, B., & Burleigh, J. (1996). Leaf color and vine size are related to yield in a phylloxera-infested vineyard. *Vitis*, *35*(4), 201–205. Retrieved from http://www.vitis-vea.de/admin/volltext/e037969.pdf

Bannari, A., Khurshid, K. S., Staenz, K., & Schwarz, J. W. (2007). A comparison of hyperspectral chlorophyll indices for wheat crop chlorophyll content estimation using laboratory reflectance measurements. In *IEEE Transactions on Geoscience and Remote Sensing* (Vol. 45, pp. 3063–3074).

Baranoski, G. V. G., & Rokne, J. G. (2005). A practical approach for estimating the red edge position of plant leaf reflectance. *International Journal of Remote Sensing*, *26*(3), 503–521. doi:10.1080/01431160512331314029

Bates, D. M., & Watts, D. G. (1988). *Nonlinear regression analysis and its applications*. Toronto, Canada: John Wiley & Sons Inc. doi:10.1002/9780470316757

Bertamini, M., & Nedunchezhian, N. (2003). Photosynthetic functioning of individual grapevine leaves (Vitis vinifera L. cv. Pinot noir) during ontogeny in the field. *Vitis*, *42*(1), 13–17.

Biochrom Ltd. (n.d.-a). *Basic UV/Visible spectrophotometry*. Cambridge, UK. Retrieved from www.biochrom.co.uk/download/72/

Biochrom Ltd. (n.d.-b). *Lightwave II & Lightwave II: User manual*. Cambridge, UK. Retrieved from http://www.biochrom.co.uk/download/86/

Blackburn, G. A. (1998). Quantifying chlorophylls and caroteniods at leaf and canopy scales: An evaluation of some hyperspectral approaches. *Remote Sensing of Environment*, *66*, 273–285. doi:10.1016/S0034-4257(98)00059-5

Blackburn, G. A. (1999). Relationships between spectral reflectance and pigment concentrations in stacks of deciduous broadleaves. *Remote Sensing of Environment*, *70*(2), 224–237. doi:10.1016/S0034-4257(99)00048-6

Blackburn, G. A. (2007). Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany*, *58*(4), 855–67. doi:10.1093/jxb/erl123

Borengasser, M., Hungate, W., & Watkinds, R. (2008). *Hyperspectral remote sensing: Principles and applications*. (Q. Weng, Ed.). Boca Raton, FL: CRC Press Taylor & Francis Group.

Bramley, R. G. V., & Hamilton, R. P. (2004). Understanding variability in winegrape production systems: 1. Within vineyard variation in quality over several vintages. *Australian Journal of Grape and Wine Research*, *10*(1), 32–45. doi:10.1111/j.1755-0238.2005.tb00277.x

British Columbia Wine Institute. (2012). *British Columbia Wine Institute Annual Report 2011-2012*. Kelowna, BC. Retrieved from http://www.winebc.org/files/Information/BCWI Annual Reports/2012AnnualReport.pdf

Broge, N. H., & Leblanc, E. (2000). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, *76*(2), 156–172. doi:10.1016/S0034-4257(00)00197-8

Broge, N. H., & Mortensen, J. V. (2002). Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. *Remote Sensing of Environment*, *81*, 45–57. doi:10.1016/S0034-4257(01)00332-7

Campbell, N. A., & Reece, J. B. (2002). *Biology*. (B. Wilbur, Ed.) (6th ed.). San Francisco, CA: Benjamin-Cummings Publishing Co.

Carsey, T. M., & Harden, J. J. (2014). Introduction. In *Monte Carlo Simulation and Resampling Methods for Social Scientists* (pp. 1–31). Thousand Oaks, CA: SAGE Publications Inc. doi:http://dx.doi.org/10.4135/9781483319605.n1

Cartelat, A., Cerovic, Z. G., Goulas, Y., Meyer, S., Lelarge, C., Prioul, J. L., … Moya, I. (2005). Optically assessed contents of leaf polyphenolics and chlorophyll as indicators of nitrogen deficiency in wheat (Triticum aestivum L.). *Field Crops Research*, *91*(1), 35–49. doi:10.1016/j.fcr.2004.05.002

Carter, G. A., & Spiering, B. A. (2002). Optical properties of intact leaves for estimating chlorophyll concentration. *Journal of Environmental Quality*, *31*(5), 1424–1432. doi:10.2134/jeq2002.1424

Cerovic, Z. G., Ghozlen, N. B., Milhade, C., Obert, M., Debuisson, S., & Le Moigne, M. (2015). Nondestructive diagnostic test for nitrogen nutrition of grapevine (*Vitis vinifera L.*) based on Dualex leaf-clip measurements in the field. *Journal of Agricultural and Food Chemistry*, *63*(14), 3669–80. doi:10.1021/acs.jafc.5b00304

Cerovic, Z. G., Masdoumier, G., Ghozlen, N. Ben, & Latouche, G. (2012). A new optical leaf-clip meter for simultaneous non-destructive assessment of leaf chlorophyll and epidermal flavonoids. *Physiologia Plantarum*, *146*(3), 251–60. doi:10.1111/j.1399-3054.2012.01639.x

Cho, M. A., & Skidmore, A. K. (2006). A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *Remote Sensing of Environment*, *101*(2), 181–193. doi:10.1016/j.rse.2005.12.011

Cho, M. A., Sobhan, I., Skidmore, A. K., & de Leeuw, J. (2008). Discriminating species using hyperspectral indices at leaf and canopy scales. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. 37, pp. 369–376). Beijing.

Chutia, J., & Borah, S. P. (2012). Water stress effects on leaf growth and chlorophyll content but not the grain yield in traditional rice (*Oryza sativa Linn.*) genotypes of Assam, India II. Protein and proline status in seedlings under PEG induced water stress. *American Journal of Plant Sciences*, *3*(7), 971–980. doi:10.4236/ajps.2012.37115

Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. In A. N. Rencz (Ed.), *Remote sensing for the earth sciences: Manual of remote sensing* (pp. 3– 52). New York, NW: John Wiley and Sons.

Clark, R. N., & Roush, T. L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, *89*(B7), 6329. doi:10.1029/JB089iB07p06329

Clevers, J. G. P. W., De Jong, S. M., Epema, G. F., van der Meer, F., Bakker, W. H., Skidmore, A. K., & Scholte, K. H. (2002). Derivation of the red edge index using the MERIS standard band setting. *International Journal of Remote Sensing*, *23*(16), 3169–3184. doi:10.1080/01431160110104647

Cochrane, M. A. (2000). Using vegetation reflectance variability for species level classification of hyperspectral data. *International Journal of Remote Sensing*, *21*(10), 2075–2087. doi:http://dx.doi.org/10.1080/01431160050021303

Coste, S., Baraloto, C., Leroy, C., Marcon, É., Renaud, A., Richardson, A. D., … Hérault, B. (2010). Assessing foliar chlorophyll contents with the SPAD-502 chlorophyll meter: a calibration test with thirteen tree species of tropical rainforest in French Guiana. *Annals of Forest Science*, *67*(6), 607–607. doi:10.1051/forest/2010020

Cui, M., Vogelmann, T. C., & Smith, W. K. (1991). Chlorophyll and light gradients in sun and shade leaves of *Spinacia oleracea*. *Plant, Cell & Environment*, *14*(5), 493–500. doi:10.1111/j.1365-3040.1991.tb01519.x

Cunha, M., Marcal, A., & Silva, L. (2010). Very early prediction of wine yield based on satellite data from VEGETATION. *International Journal of Remote Sensing*, *31*(12), 3125–3142. doi:10.1080/01431160903154382

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, *30*(3), 271–278. doi:10.1016/0034-4257(89)90069-2

Curran, P. J., Dungan, J. L., Macler, B. A., & Plummer, S. E. (1991). The effect of a red leaf pigment on the relationship between red edge and chlorophyll concentration. *Remote Sensing of Environment*, *35*(1), 69–76. doi:10.1016/0034-4257(91)90066-F

Curran, P. J., Dungan, J. L., & Peterson, D. L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry. Testing the Kokaly and Clark methodologies. *Remote Sensing of Environment*, *76*, 349–359.

Danehower, C. (2010). *Essential wines and wineries of the Pacific Northwest*. Portland, Oregon: Timber Press.

Datt, B. (1998). Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+b, and total carotenoid content in Eucalyptus leaves. *Remote Sensing of Environment*, *66*(2), 111–121. doi:10.1016/S0034-4257(98)00046-7

Datt, B. (1999). Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *International Journal of Remote Sensing*, *20*(14), 2741–2759. doi:10.1080/014311699211778

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B., & McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, *74*(2), 229–239. doi:10.1016/S0034-4257(00)00113-9

Desclée, B., Bogaert, P., & Defourny, P. (2006). Forest change detection by statistical object-based method. *Remote Sensing of Environment*, *102*(1-2), 1–11. doi:10.1016/j.rse.2006.01.013

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Elvidge, C. D., & Chen, Z. (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sensing of Environment*, *54*(1), 38–48. doi:10.1016/0034-4257(95)00132-K

Fadock, M. (2011). *Non-destructive VIS-NIR reflectance spectrometry for red wine grape analysis* (Master's thesis). University of Guelph. Retrieved from https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/2827/Michael Fadock, Masters Thesis, Submitted.pdf?sequence=11

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, UK: SAGE Publications Inc.

Fiorillo, E., Crisci, A., De Filippis, T., Di Gennaro, S. F., Di Blasi, S., Matese, A., … Genesio, L. (2012). Airborne high-resolution images for grape classification: Changes in correlation between technological and late maturity in a Sangiovese vineyard in Central Italy. *Australian Journal of Grape and Wine Research*, *18*(1), 80–90. doi:10.1111/j.1755-0238.2011.00174.x

Gao, J. (2006). *Canopy chlorophyll estimation with hyperspectral remote sensing* (Doctoral dissertation). Kansas State University, USA. Retrieved from https://krex.kstate.edu/dspace/bitstream/handle/2097/252/JinGao2006.pdf?sequence=1

Gausman, H. W. (1984). Evaluation of factors causing reflectance differences between sun and shade Leaves. *Remote Sensing of Environment*, *15*(1984), 177–181. doi:10.1016/0034-4257(84)90045-2

Girma, A., Skidmore, A. K., de Bie, C. A. J. M., Bongers, F., & Schlerf, M. (2013). Photosynthetic bark: Use of chlorophyll absorption continuum index to estimate *Boswellia papyrifera* bark chlorophyll content. *International Journal of Applied Earth Observation and Geoinformation*, *23*(1), 71–80. doi:10.1016/j.jag.2012.10.013

Gitelson, A. (2012). Nondestructive estimation of foliar pigment (chlorophylls, carotenoids, and anthocyanins) contents: Evaluating a semianalytical three-band model. In P. S. Thenkabail, J. G. Lyon, & A. R. Huete (Eds.), *Hyperspectral remote sensing of vegetation* (pp. 141–166). Boka Raton, FL: CRC Press Taylor & Francis Group.

Gitelson, A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, *160*(3), 271–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12749084

Gitelson, A., Keydan, G. P., & Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Research Letters*, *33*(11), L11402. doi:10.1029/2006GL026457

Gitelson, A., & Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, *18*(12), 2691–2697. doi:10.1080/014311697217558

Gitelson, A., Merzlyak, M. N., & Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *Journal of Plant Physiology*, *148*, 501–508. doi:10.1016/S0176-1617(96)80285-9

Gitelson, A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, *32*, 1–4. doi:10.1029/2005GL022688

Google Inc. (2016). [Venturi-Schulze Vineyard]. Retrieved February 12, 2016, from https://www.google.ca/maps/@48.7178364,123.6200752,4181m/data=!3m1!1e3?hl=en

Grossman, Y. L., Ustin, S. L., Jacquemoud, S., Sanderson, E. W., Schmuck, G., & Verdeboutt, J. (1996). Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. *Remote Sensing of Environment*, *56*, 182–193.

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. doi:10.1016/j.rse.2003.12.013

Haboudane, D., Tremblay, N., Miller, J. R., & Vigneault, P. (2008). Remote estimation of crop chlorophyll content using spectral indices derived from hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(2), 423–436. doi:10.1109/TGRS.2007.904836

Hall, A., Lamb, D. W., Holzapfel, B., & Louis, J. (2002). Optical remote sensing applications in viticulture—A review. *Australian Journal of Grape and Wine Research*, *8*(1), 36–47. doi:10.1111/j.1755-0238.2002.tb00209.x

Hall, A., Louis, J., & Lamb, D. W. (2003). Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images. *Computers & Geosciences*, *29*(7), 813–822. doi:10.1016/S0098-3004(03)00082-7

Harrell, F. J. (2015). *Regression Modeling Strategies*. Nashville, TN. Retrieved from Vanderbit University School of Medicine: http://biostat.mc.vanderbilt.edu/wiki/pub/Main/RmS/rms.pdf

Hayes, A. F., & Cai, L. (2007). Using heteroscedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*(4), 709–722.

Hiscox, J. D., & Israelstam, G. F. (1979). A method for the extraction of chlorophyll from leaf tissue without maceration. *Canadian Journal of Botany*, *57*(12), 1332–1334. doi:10.1139/b79-163

Huang, Z., Turner, B. J., Dury, S. J., Wallis, I. R., & Foley, W. J. (2004). Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sensing of Environment*, *93*, 18–29. doi:10.1016/j.rse.2004.06.008

Hynes, G. (2011). *Island wineries of British Columbia*. (M. Horsdal, Ed.). Surry, BC: TouchWood Editions.

Ibrahim, M. H., & Jaafar, H. Z. E. (2013). Relationship between extractable chlorophyll content and SPAD values in three varieties of Kacip Fatimah under greenhouse conditions. *Journal of Plant Nutrition*, *36*(9), 1366–1372. doi:10.1080/01904167.2013.792836

Jackson, R. D., & Huete, A. R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*, *11*(3-4), 185–200. doi:10.1016/S0167-5877(05)80004-2

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., … Ustin, S. L. (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, *113*, S56–S66. doi:10.1016/j.rse.2008.01.026

Jago, R. A., Cutler, M. E. J., & Curran, P. J. (1999). Estimating canopy chlorophyll concentration from field and airborne spectra. *Remote Sensing of Environment*, *68*(3), 217–224. doi:10.1016/S0034-4257(98)00113-8

Jensen, J. R. (2007). *Remote sensing of the environment: An Earth resource perspective* (2nd ed.). New Jersey: Prentice Hall Inc.

Johnson, L., Nemani, R., Hornbuckle, J., Bastiaanssen, W., Thoreson, B., Tisseyre, B., & Pierce, L. (2012). Remote sensing for viticultural research and production. In P. H. Dougherty (Ed.), *The Geography of Wine* (pp. 209–226). Netherlands: Springer. doi:10.1007/978-94-007-0464-0

Johnson, L., Roczen, D. E., Youkhana, S. K., Nemani, R. R., & Bosch, D. F. (2003). Mapping vineyard leaf area with multispectral satellite imagery. *Computers and Electronics in Agriculture*, *38*(1), 33–44. doi:10.1016/S0168-1699(02)00106-0

Kırca, A., Yemiş, O., & Özkan, M. (2006). Chlorophyll and colour changes in grapevine leaves preserved by passive modification. *European Food Research and Technology*, *223*(3), 387–393. doi:10.1007/s00217-005-0216-6

Kochubey, S. M., & Kazantsev, T. A. (2007). Changes in the first derivatives of leaf reflectance spectra of various plants induced by variations of chlorophyll content. *Journal of Plant Physiology*, *164*(12), 1648–55. doi:10.1016/j.jplph.2006.11.007

Kokaly, R. F. (2001). Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment*, *75*, 153–161.

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*, 267–287.

Kokaly, R. F., Despain, D. G., Clark, R. N., & Livo, K. E. (2003). Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. *Remote Sensing of Environment*, *84*, 437–456.

Koundouras, S., Tsialtas, I. T., Zioziou, E., & Nikolaou, N. (2008). Rootstock effects on the adaptive strategies of grapevine (*Vitis vinifera* L. cv. Cabernet-Sauvignon) under contrasting water status: Leaf physiological and structural responses. *Agriculture, Ecosystems and Environment*, *128*(1-2), 86–96. doi:10.1016/j.agee.2008.05.006

Lamb, D. W., Steyn-Ross, M., Schaare, P., Hanna, M. M., Silvester, W., & Steyn-Ross, A. (2002). Estimating leaf nitrogen concentration in ryegrass (*Lolium spp.*) pasture using the chlorophyll red-edge: Theoretical modelling and experimental observations. *International Journal of Remote Sensing*, *23*(18), 3619–3648. doi:10.1080/01431160110114529

Lamb, D. W., Weedon, M. M., & Bramley, R. G. V. (2004). Using remote sensing to predict grape phenolics and colour at harvest in a Cabernet Sauvignon vineyard: Timing observations against vine phenology and optimising image resolution. *Australian Journal Of Grape And Wine Research*, *10*, 46–54.

Lashbrooke, J. G., Young, P. R., Strever, A. E., Stander, C., & Vivier, M. A. (2010). The development of a method for the extraction of carotenoids and chlorophylls from grapevine leaves and berries for HPLC profiling. *Australian Journal of Grape and*

Wine Research, *16*(2), 349–360. doi:10.1111/j.1755-0238.2010.00097.x

Lausch, A., Pause, M., Schmidt, A., Salbach, C., Gwillym-Margianto, S., & Merbach, I. (2013). Temporal hyperspectral monitoring of chlorophyll, LAI, and water content of barley during a growing season. *Canadian Journal of Remote Sensing*, *39*(3), 191–207. doi:10.5589/m13-028

Lawrence, R., & Labus, M. (2003). Early Detection of Douglas-Fir Beetle Infestation with Subcanopy Resolution Hyperspectral Imagery. *Western Journal of Applied Forestry*, *18*(3), 1–5.

Le Maire, G., François, C., & Dufrêne, E. (2004). Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment*, *89*(1), 1–28. doi:10.1016/j.rse.2003.09.004

Legendre, P., & Legendre, L. (1998). *Numerical ecology* (2nd ed.). Amseterdam: Elsevier Inc.

Li, F., Miao, Y., Hennig, S. D., Gnyp, M. L., Chen, X., Jia, L., & Bareth, G. (2010). Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages. *Precision Agriculture*, *11*(4), 335–357. doi:10.1007/s11119-010-9165-6

Lichtenthaler, H. K., Gitelson, A., & Lang, M. (1996). Non-destructive determination of chlorophyll content of leaves of a green and an aurea mutant of tobacco by reflectance measurements. *Journal of Plant Physiology*, *148*(3-4), 483–493. doi:10.1016/S0176-1617(96)80283-5

Lichtfouse, E., Navarrete, M., Debaeke, P., Souchere, V., Alberola, C., & Menassieu, J. (2009). Agronomy for sustainable agriculture. A review. *Agronomy for Sustainable Development*, *29*(1), 1–6. doi:10.1051/agro:2008054

Lin, C., Popescu, S. C., Huang, S. C., Chang, P. T., & Wen, H. L. (2015). A novel reflectance-based model for evaluating chlorophyll concentration of fresh and water-stressed leaves. *Biogeosciences*, *12*, 49–66. doi:10.5194/bgd-10-17893-2013

Ling, Q., Huang, W., & Jarvis, P. (2011). Use of a SPAD-502 meter to measure leaf chlorophyll concentration in *Arabidopsis thaliana*. *Photosynthesis Research*, *107*(2), 209–214. doi:10.1007/s11120-010-9606-0

Luoheng, H., & Rundquist, D. C. (1997). Comparison of NIR/RED ratio and first derivative of reflectance in estimating algal-chlorophyll concentration: A case study in a turbid reservoir. *Remote Sensing of Environment*, *62*(3), 253–261. doi:10.1016/S0034-4257(97)00106-5

Maccioni, A., Agati, G., & Mazzinghi, P. (2001). New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *Journal of Photochemistry and Photobiology B: Biology*, *61*(1-2), 52–61. doi:10.1016/S1011-1344(01)00145-2

Main, R., Cho, M. A., Mathieu, R., O'Kennedy, M. M., Ramoelo, A., & Koch, S. (2011). An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), 751–761. doi:10.1016/j.isprsjprs.2011.08.001

Malenovský, Z., Homolová, L., Zurita-Milla, R., Lukeš, P., Kaplan, V., Hanuš, J., … Schaepman, M. E. (2013). Retrieval of spruce leaf chlorophyll content from airborne image data using continuum removal and radiative transfer. *Remote Sensing of Environment*, *131*, 85–102. doi:10.1016/j.rse.2012.12.015

Malenovský, Z., Ufer, C., Lhotáková, Z., Clevers, J. G. P. W., Schaepman, M. E., Albrechtova, J., … Z Malenovský Z Lhotáková, J G P W Clevers, M E Schaepman, J Albrechtová & P Cudlín, C. U. (2006). A new hyperspectral index for chlorophyll estimation of a forest canopy: Area under curve normalised to maximal band depth between 650-725 nm. *EARSeL eProceedings, 5*(2), 161–172.

Marenco, R. A., Antezana-Vera, S. A., & Nascimento, H. C. S. (2009). Relationship between specific leaf area, leaf thickness, leaf water content and SPAD-502 readings in six Amazonian tree species. *Photosynthetica*, *47*(2), 184–190. doi:10.1007/s11099-009-0031-6

Markwell, J., Osterman, J. C., & Mitchell, J. L. (1995). Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynthesis Research*, *46*(3), 467–472. doi:10.1007/BF00032301

Martín, P. M., Zarco-Tejada, P. J., Gonzalez, M., & Berjón, A. (2007). Using hyperspectral remote sensing to map grape quality in 'Tempranillo' vineyards affected by iron deficiency chlorosis. *Vitis*, *46*(1), 7–14.

Matese, A., Toscano, P., Di Gennaro, S., Genesio, L., Vaccari, F., Primicerio, J., … Gioli, B. (2015). Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote Sensing*, *7*(3), 2971–2990. doi:10.3390/rs70302971

Mathews, A., & Jensen, J. (2013). Visualizing and quantifying vineyard canopy LAI using an unmanned aerial vehicle (UAV) collected high density structure from motion point cloud. *Remote Sensing*, *5*(5), 2164–2183. doi:10.3390/rs5052164

May, D., & Sharpe, A. (1997). *The Everything Wine Book*. Holbrook, MA: Adams Media Corporation.

McCoy, R. (2005). *Field methods in remote sensing*. New York, NY: The Guilford Press.

McKillup, S. (2012). *Statistics Explained* (2nd ed.). Cambridge, England: Cambridge University Press.

Meeuwig, J. J., & Peters, R. H. (1996). Circumventing phosphorus in lake management: A comparison of chlorophyll a predictions from land-use and phosphorus-loading models. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*(8), 1795–1806. doi:10.1139/cjfas-53-8-1795

Mettler-Toledo. (2012). *Classic Light Balances: Operating Instructions*. Mettler-Toledo AG, Laboratory & Weighing Technologies. Greifensee, Switzerland.

Milton, E. J., Schaepman, M. E., Anderson, K., Kneubühler, M., & Fox, N. (2009). Progress in field spectroscopy. *Remote Sensing of Environment*, *113*, S92–S109. doi:10.1016/j.rse.2007.08.001

Monje, O. A., & Bugbee, B. (1992). Inherent limitations of nondestructive chlorophyll meters: A comparison of two types of meters. *HortScience*, *27*(1), 69–71.

Murphy, R. J., Tolhurst, T. J., Chapman, M. G., & Underwood, A. J. (2005). Estimation of surface chlorophyll-a on an emersed mudflat using field spectrometry: Accuracy of ratios and derivative-based approaches. *International Journal of Remote Sensing*, *26*(9), 1835–1859. doi:10.1080/01431160512331326530

Mutanga, O., & Skidmore, A. K. (2004). Hyperspectral band depth analysis for a better estimation of grass biomass (*Cenchrus ciliaris*) measured under controlled laboratory conditions. *International Journal of Applied Earth Observation and Geoinformation*, *5*(2), 87–96. doi:10.1016/j.jag.2004.01.001

Mutanga, O., & Skidmore, A. K. (2007). Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry and Remote Sensing*, *62*(1), 34–42. doi:10.1016/j.isprsjprs.2007.02.001

Mutanga, O., Skidmore, A. K., Kumar, L., & Ferwerda, J. (2005). Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *International Journal of Remote Sensing*, *26*(6), 1093–1108. doi:10.1080/01431160512331326738

Mutanga, O., Skidmore, A. K., & Prins, H. H. T. (2004). Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote Sensing of Environment*, *89*(3), 393–408. doi:10.1016/j.rse.2003.11.001

Nascimento, H. C. S., & Marenco, R. A. (2010). SPAD-502 readings in response to photon fluence in leaves with different chlorophyll content. *Revista Ceres*, *57*, 614–620. doi:10.1590/S0034-737X2010000500008

Nauš, J., Prokopová, J., Řebíček, J., & Špundová, M. (2010). SPAD chlorophyll meter reading can be pronouncedly affected by chloroplast movement. *Photosynthesis Research*, *105*(3), 265–271. doi:10.1007/s11120-010-9587-z

Netto, A., Campostrini, E., Oliveira, J., & Bressansmith, R. (2005). Photosynthetic pigments, nitrogen, chlorophyll fluorescence and SPAD-502 readings in coffee leaves. *Scientia Horticulturae*, *104*(2), 199–209. doi:10.1016/j.scienta.2004.08.013

Niemann, K. O., Quinn, G. S., Stephen, R., Visintini, F., & Parton, D. (2015). Hyperspectral remote sensing of mountain pine beetle with an emphasis on previsual assessment. *Canadian Journal of Remote Sensing*, *41*, 191–202. doi:10.1080/07038992.2015.1065707

Novoa, S., Chust, G., Valencia, V., Froidefond, J.-M., & Morichon, D. (2011). Estimation of chlorophyll-a concentration in waters over the continental shelf of the Bay of Biscay: a comparison of remote sensing algorithms. *International Journal of Remote Sensing*, *32*(23), 8349–8371. doi:10.1080/01431161.2010.540588

Ozelkan, E., Karaman, M., Candar, S., Coskun, Z., & Ormeci, C. (2015). Investigation of grapevine photosynthesis using hyperspectral techniques and development of hyperspectral band ratio indices sensitive to photosynthesis. *Journal of Environmental Biology*, *36*, 91–100.

Pannucci, C., & Wilkins, E. (2010). Identifying and avoiding bias in research. *Plastic and Reconstructive Surgery*, *126*(2), 619–625. doi:10.1097/PRS.0b013e3181de24bc.Identifying

Parton, D., Quinn, G. S., Visintini, F., & Niemann, K. O. (2012). A temporal investigation of leaf spectroscopy for grape variety discrimination. In *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2012 4th Workshop on*. Shanghai, China: IEEE. doi:10.1109/WHISPERS.2012.6874233

Pu, R., Ge, S., Kelly, N. M., & Gong, P. (2003). Spectral absorption features as indicators of water status in coast live oak (*Quercus agrifolia*) leaves. *International Journal of Remote Sensing*, *24*(9), 1799–1810. doi:10.1080/01431160210155965

Qin, J., Rundquist, D., Gitelson, A., Tan, Z., & Steele, M. R. (2011). A non-linear model of nondestructive estimation of anthocyanin content in grapevine leaves with visible/red-infrared hyperspectral. In *International Federation For Information Processing* (pp. 47–62).

Quinn, G. S. (2010). *Towards an operational root disease mapping methodology through lidar integrated imaging spectroscopy* (Master's thesis). University of Victoria. Retrieved from https://dspace.library.uvic.ca/bitstream/handle/1828/3605/Quinn_MSc.pdf?sequence=1&isAllowed=y

Quinn, G. S., Visintini, F., & Niemann, K. O. (2012). Considering the implications of species on pigment estimation from leaf spectroscopy. In *4th Workshop on Hyperspectral Image and Signal Processing (WHISPERS)* (pp. 12–15).

Ranjit, K. P. (2006). *Some Methods of Detection of Outliers in Linear Regression Model*. Indian Agriculture Statistics Research Group. New Dheli.

Rawlings, J. O., Pantula, Sastry, G., & Dickey, D. A. (1998). *Applied regression analysis: A research tool* (2nd ed.). New York, NY: Springer. Retrieved from http://web.nchu.edu.tw/~numerical/course1012/ra/Applied_Regression_Analysis_A_Research_Tool.pdf

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

Richardson, A. D., Duigan, S. P., & Berlyn, G. P. (2002). An evaluation of noninvasive methods to estimate foliar chlorophyll content. *New Phytologist*, *153*(1), 185–194. doi:10.1046/j.0028-646X.2001.00289.x

Roberts, D. A., Roth, K. L., & Perroy, R. L. (2011). Hyperspectral vegetation indices. In P. S. Thenkabail, J. G. Lyon, & A. R. Huete (Eds.), *Hyperspectral remote sensing of vegetation* (pp. 309–327). Boca Raton, FL: CRC Press Taylor & Francis Group. doi:10.1201/b11222-20

Robinson, J. (Ed.). (1994). *Oxford Companion to Wine* (1st ed.). Oxford, UK: Oxford University Press.

Rock, B. N., Hoshizaki, T., & Miller, J. R. (1988). Comparison of in situ and airborne spectral measurements of the blue shift associated with forest decline. *Remote Sensing of Environment*, *24*, 109–127.

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.

Rosellpolo, J., Sanz, R., Llorens, J., Arnó, J., Escola, A., Ribes-Dasi, M., … Palacin, J. (2009). A tractor-mounted scanning LIDAR for the non-destructive measurement of vegetative volume and surface area of tree-row plantations: A comparison with conventional destructive measurements. *Biosystems Engineering*, *102*(2), 128–134. doi:10.1016/j.biosystemseng.2008.10.009

Rouse, J. W., Hass, R. H., Deering, D. W., & Schell, J. A. (1973). *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation* (Type II Report). Goddard Space Flight Center/NASA. Greenbelt, Maryland.

Ruiz-Espinoza, F. H., Murillo-Amador, B., García-Hernández, J. L., Fenech-Larios, L., Rueda-Puente, E. O., Troyo-Diéguez, E., … Beltrán-Morales, A. (2010). Field evaluation of the relationship between chlorophyll content in basil leaves and a portable chlorophyll meter (SPAD-502) readings. *Journal of Plant Nutrition*, *33*(3), 423–438. doi:10.1080/01904160903470463

Saari, H., Pellikka, I., Pesonen, L., Tuominen, S., Heikkilä, J., Holmlund, C., … Antila, T. (2011). Unmanned aerial vehicle (UAV) operated spectral camera system for forest and agriculture applications. In C. M. U. Neale, A. Maltese, & K. Richter (Eds.), *Proc. SPIE, Remote Sensing for Agriculture, Ecosystems, and Hydrology XIII* (Vol. 8174). doi:10.1117/12.897585

Sanches, I. D., Filho, C. R., & Kokaly, R. F. (2014). Spectroscopic remote sensing of plant stress at leaf and canopy levels using the chlorophyll 680nm absorption feature with continuum removal. *ISPRS Journal of Photogrammetry and Remote Sensing*, *97*, 111–122. doi:10.1016/j.isprsjprs.2014.08.015

Sanches, I. D., & Filho, D. S. (2014). Continuum removal analysis of the chlorophyll absorption feature to detect plant stress induced by liquid hydrocarbon contamination. In *IGARSS 2014*. Quebec, Canada. Retrieved from ftp://ftp.legos.obs-mip.fr/pub/tmp3m/IGARSS2014/abstracts/1407.pdf

Schlemmer, M. R., Francis, D. D., Shananhan, J. F., & Schepers, J. S. (2005). Remotely measuring chlorophyll content in corn leaves with differing nitrogen levels and relative water content. *Agronomy Journal*, *97*, 106–112.

Schlerf, M., Atzberger, C., Udelhoven, T., Jarmer, T., Mader, S., Werner, W., & Hill, J. (2003). Spectrometric estimation of leaf pigments in Norway spruce needles using band-depth analysis, partial least-square regression and inversion of a conifer leaf model. In *3rd EARSeL Workshop on Imaging Spectroscopy*. Herrsching.

Schmidt, K. S., & Skidmore, A. K. (2010). Exploring spectral discrimination of grass species in African rangelands. *International Journal of Remote Sensing*, *22*(17), 3421–3434. doi:10.1080/01431160152609245

Serrano, L., González-Flor, C., & Gorchs, G. (2012). Assessment of grape yield and composition using the reflectance based water index in Mediterranean rainfed vineyards. *Remote Sensing of Environment*, *118*, 249–258. doi:10.1016/j.rse.2011.11.021

Serrano, L., Peñuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2-3), 355–364. doi:10.1016/S0034-4257(02)00011-1

Sims, D., & 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. doi:10.1016/S0034-4257(02)00010-X

Smith, K. L., Steven, M. D., & Colls, J. J. (2004). Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks. *Remote Sensing of Environment*, *92*(2), 207–217. doi:10.1016/j.rse.2004.06.002

Sommers, B. (2008). *The geography of wine*. New York, New York: Penguin Group.

Spectrum Technologies Inc. (2009). *SPAD 502 Plus chlorophyll meter product manual*. Aurora, IL. Retrieved from http://www.specmeters.com/assets/1/22/2900P_SPAD_502.pdf

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.

Stagakis, S., Markos, N., Sykioti, O., & Kyparissis, A. (2010). Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a *Phlomis fruticosa* Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sensing of Environment*, *114*(5), 977–994. doi:10.1016/j.rse.2009.12.006

Staid, M., & Staid, M. (2009). VineView: Scientific Aerial Imaging. Retrieved December 7, 2015, from http://www.vineview.com

Statistics Canada. (2011). *Thematic maps from the Census of Agriculture: Grape Area, 2011*. Ottawa, Ontario. . Retrieved August 2015, from http://www.statcan.gc.ca/pub/95-634-x/2014001/14067-eng.htm.

Statistics Canada. (2014). *Control and sale of alcoholic beverages, for the year ending March 31, 2013*. Ottawa, Ontario.

Statistics Canada. (2015). *Control and sale of alcoholic beverages, for the year ending March 31, 2014*. Ottawa, Canada.

Steele, M. R. (2007). *Non-destructive estimation of leaf pigments and monitoring phenology of grapevines* (Master's thesis). University of Nebraska-Lincoln.

Steele, M. R., Gitelson, A., & Rundquist, D. (2008a). Nondestructive estimation of leaf chlorophyll content in grapes. *American Journal of Enology and Viticulture*, *59*(3), 299–305.

Steele, M. R., Gitelson, A., & Rundquist, D. C. (2008b). A comparison of two techniques for nondestructive measurement of chlorophyll content in grapevine leaves. *Agronomy Journal*, *100*(3), 779–782. doi:10.2134/agronj2007.0254N

Strever, A. E. (2003). *A study of within-vineyard variability with conventional and remote sensing technology* (Master's thesis). Stellenbosch University. Retrieved from http://scholar.sun.ac.za/handle/10019.1/53304

Strever, A. E. (2012). *Non-destructive assessment of leaf composition as related to growth of the grapevine (Vitis vinifera L. cv. Shiraz)* (Doctoral dissertation). Stellenbosch University. Retrieved from http://scholar.sun.ac.za/handle/10019.1/19963

Stroppiana, D., Fava, F., Boschetti, M., & Brivio, P. A. (2011). Estimation of nitrogen content in crops and pastures using hyperspectral vegetation indices. In P. S. Thenkabail, J. G. Lyon, & A. R. Huete (Eds.), *Hyperspectral remote sensing of vegetation* (pp. 245–262). Boca Raton, FL: CRC Press Taylor & Francis Group.

Sumanta, N., Haque, C. I., Nishika, J., & Suprakash, R. (2014). Spectrophotometric analysis of chlorophylls and carotenoids from commonly grown fern species by using various extracting solvents. *Research Journal of Chemical Sciences*, *4*(9), 63–69.

Taskos, D. G., Koundouras, S., Stamatiadis, S., Zioziou, E., Nikolaou, N., Karakioulakis, K., & Theodorou, N. (2014). Using active canopy sensors and chlorophyll meters to estimate grapevine nitrogen status and productivity. *Precision Agriculture*, *16*(1), 77–98. doi:10.1007/s11119-014-9363-8

Thenkabail, P. S., Enclona, E. a., Ashton, M. S., & Van Der Meer, B. (2004). Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sensing of Environment*, *91*(3-4), 354–376. doi:10.1016/j.rse.2004.03.013

Thenkabail, P. S., Lyon, J. G., & Huete, A. (2011). Hyperspectral remote sensing of vegetation and agricultural crops: knowledge gain and knowledge gap after 40 years of research. In P. S. Thenkabail, J. G. Lyon, & A. R. Huete (Eds.), *Hyperspectral remote sensing of vegetation* (pp. 663–688). Boca Raton, FL: CRC Press Taylor & Francis Group.

Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, *71*(2), 158–182. doi:10.1016/S0034-4257(99)00067-X

Torres, P. B., Chow, F., Furlan, C. M., Mandelli, F., Mercadante, A., & dos Santos, D. Y. A. C. (2014). Standardization of a protocol to extract and analyze chlorophyll *a* and carotenoids in *Gracilaria tenuistipitata* var. *liui.* Zhang and Xia (rhodophyta). *Brazilian Journal of Oceanography*, *62*(1), 57–63.

Townend, J. (2002). *Practical Statistics for environmental and Biological Scientists*. West Sussex, England: John Wiley & Sons Ltd.

Tsai, F., & Philpot, W. (1998). Derivative Analysis of Hyperspectral Data. *Remote Sensing of Environment*, *66*, 41–51.

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

Uddling, J., Gelang-Alfredsson, J., Piikki, K., & Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research*, *91*(1), 37–46. doi:10.1007/s11120-006-9077-5

Uno, Y. (2003). *Application of machine learning methods and airborne hyperspectral remote sensing for crop yield estimation* (Master's thesis). McGill University. Retrieved from http://webpages.mcgill.ca/staff/deptshare/FAES/066-Bioresource/Theses/theses/298YojiUno2003/298YojiUno2003.pdf

Ustin, S. L., Gitelson, A., Jacquemoud, S., Schaepman, M., Asner, G. P., Gamon, J. A., & Zarco-Tejada, P. J. (2009). Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment*, *113*, S67–S77. doi:10.1016/j.rse.2008.10.019

Ustin, S. L., Roberts, D. A., Gamon, J. A., Asner, G. P., & Green, R. O. (2004). Using imaging spectroscopy to study ecosystem processes and properties. *BioScience*, *54*(6), 523. doi:10.1641/0006-3568(2004)054[0523:UISTSE]2.0.CO;2

van Belle, G. (2002). *Statistical Rules of Thumb*. New York: John Wiley & Sons Inc.

van der Meer, F. (2004). Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, *5*(1), 55–68. doi:10.1016/j.jag.2003.09.001

Viau, A. A., Jang, J., Payan, V., & Devost, A. (2005). The use of airborne LIDAR and multispectral sensors for orchard trees inventory and characterization. In *Information and Technology for Sustainable Fruit and Vegetable Production, FRUTIC 05* (pp. 12–16). Montpellier, France.

Vincent, R. (1997). *Fundamentals of Geological and Environmental Remote Sensing*. Upper Saddle River, NJ: Prentice Hall Press, Inc.

Visintini, F. (2010). *Assessment of two spectral reflectance techniques for the estimation of fuel moisture content, equivalent water thickness, and specific leaf weight in Douglas-fir (Pseudotsuga menziesii (Mirb) Franco) needles* (Master's thesis). University of Victoria. Retrieved from http://hdl.handle.net/1828/3369

Vogelmann, T. C. (1993). Plant tissue optics. *Annual Review of Plant Physiology and Plant Molecular Biology*, *44*, 231–351. Retrieved from http://www.annualreviews.org/doi/abs/10.1146/annurev.pp.44.060193.001311?journalCode=arplant.2

Vogt, P. W. (1993). *Dictionary of statistics and methodology*. Newbury Park, California: SAGE Publications Inc.

Wellburn, A. R. (1994). The spectral determination of chlorophylls *a* and *b*, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of Plant Physiology*, *144*, 307–313.

Williams, G. (2008). *Data Mining with Rattle and R*. (R. Gentleman, K. Hornik, & G. Parmigiani, Eds.) *Applied Spatial Data Analysis with R*. Springer. doi:10.1007/978-0-387-78171-6

Wu, C., Niu, Z., Tang, Q., & Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology*, *148*, 1230–1241. doi:10.1016/j.agrformet.2008.03.005

Xiang, H., & Tian, L. (2011). Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosystems Engineering*, *108*(2), 174–190. doi:10.1016/j.biosystemseng.2010.11.003

Xue, L., & Yang, L. (2009). Deriving leaf chlorophyll content of green-leafy vegetables from hyperspectral reflectance. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(1), 97–106. doi:10.1016/j.isprsjprs.2008.06.002

Yoder, B. J., & Pettigrew-Crosby, R. E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, *53*, 199–211.

Zarco-Tejada, P. J., Berjón, A., Lopezlozano, R., Miller, J., Martin, P., Cachorro, V., … Defrutos, A. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, *99*(3), 271–287. doi:10.1016/j.rse.2005.09.002

Zarco-Tejada, P. J., Pushnik, J. C., Dobrowski, S., & Ustin, S. L. (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. doi:10.1016/S0034-4257(02)00113-X

Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture - A worldwide overview. *Computers and Electronics in Agriculture*, *36*, 113–132.

Zhang, Y. (2011). Forest leaf chlorophyll study using hyperspectral remote sensing. In P. S. Thenkabail, J. G. Lyon, & A. R. Huete (Eds.), *Hyperspectral remote sensing of vegetation* (pp. 167–186). Boca Raton, FL: CRC Press Taylor & Francis Group.

Zhang, Y., Chen, J. M., Miller, J., & Noland, T. (2008). Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sensing of Environment*, *112*(7), 3234–3247. doi:10.1016/j.rse.2008.04.005

Zhang, Y., Chen, J. M., & Thomas, S. C. (2007). Retrieving seasonal variation in chlorophyll content of overstory and understory sugar maple leaves from leaf-level hyperspectral data. *Canadian Journal of Remote Sensing*, *33*(5), 406–415. doi:10.5589/m07-037