

**USING HYPERSPECTRAL IMAGING TO QUANTIFY CADMIUM  
STRESS AND ESTIMATE CONCENTRATION IN PLANT LEAVES**

by

**Maria Paula Zea Rojas**

**A Thesis**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Master of Science**



Department of Horticulture

West Lafayette, Indiana

August 2020

**THE PURDUE UNIVERSITY GRADUATE SCHOOL  
STATEMENT OF COMMITTEE APPROVAL**

**Dr. Lori A. Hoagland, Chair**

School of Agriculture

**Dr. Yang Yang**

School of Agriculture

**Dr. Linda Lee**

School of Agriculture

**Dr. Krishna Nemali**

School of Agriculture

**Approved by:**

Dr. Aaron J. Patton

Dedicated to God who makes everything possible and to my parents who have helped me  
throughout my life

## ACKNOWLEDGMENTS

As a biologist, I envisioned myself as a scientist looking for the best for my future. So, I applied to the Cacao for Peace program and Purdue University, finding excellent personnel.

I want to express my sincerest thanks to my thesis director and my advisor, Dr. Lori Hoagland, for trusting me and helping guide my project considering my academic strengths. Her constant teaching and advice in the development of this research contributed to my training as a researcher. I want to thank all the members of the committee: Dr. Linda Lee, for all her help in the chemistry section of the study; Dr. Yang, Yang, for his teachings and dedication in hyperspectral imaging and the patience that he had during the entire project; Dr. Krishna Nemali, for every suggestion she gave me to contribute to this research and improve my knowledge.

I want to express my gratitude to all the postdocs who were part of this project. I especially want to thank Dr. Augusto de Souza for his constant help with machine learning and deep learning, and for every contribution he made to make this project possible. To all my teachers who supported me and helped me understand and analyze phenotypic images and organize my idea to translate it into this thesis, thank you. Special thanks to the greenhouse members who helped me with my plants during this pandemic, and throughout the journey and to Chris Hoagland who was aware of my project at CEPF and provided guidance.

Thanks to the USAID / USDA Cocoa for Peace program for allowing me to be a member of the program's research team and providing the resources together with Purdue University to carry out this project. Thank you to all my laboratory colleagues for the contributions to each work meeting. They allowed me to perfect my work with their unconditional support in the laboratory and their emotional help in times of stress. Finally, I want to express my deepest thanks to my parents and my partner in crime (Alex) for their unconditional supports.

# TABLE OF CONTENTS

LIST OF TABLES .....	7
LIST OF FIGURES .....	8
ABSTRACT .....	10
CHAPTER 1. INTRODUCTION .....	12
1.1 Cacao history and production .....	13
1.2 Cacao production in Colombia .....	14
1.3 Cadmium in soils .....	18
1.4 Using leafy greens as indicator crops .....	19
1.5 Using hyperspectral imaging to detect Cd stress and quantify Cd accumulation in edible plant tissues.....	20
CHAPTER 2. USING HYPERSPECTRAL IMAGING TO DETECT CADMIUM STRESS IN KALE AND BASIL AND DETERMINE WETHER BIOCHAR CAN REDUCE UPTAKE INTO EDIBLE PLANT TISSUE. ....	22
2.1 Abstract.....	22
2.2 Introduction.....	22
2.3 Materials and Methods.....	28
2.3.1 Experimental design.....	28
2.3.2 Manual plant measurements .....	29
2.3.3 RGB and hyperspectral imaging.....	30
2.3.4 Spectral processing and vegetative indices.....	30
2.3.5 Colormaps and reflectance graphs .....	32
2.3.6 Elemental concentrations in aboveground plant biomass.....	32
2.3.7 Statistical analysis.....	32
2.4 Results.....	33
2.4.1 Influence of soil Cd concentration on the development, and dry weight, elemental concentrations and SPAD readings of basil and kale .....	33
2.4.2 Influence biochar addition on the growth, dry weight, elemental concentrations and SPAD readings of basil and kale plants subject to one of four soil Cd concentrations.....	42

2.4.3	Effect of soil Cd concentrations on plant physiological indices generating using HSI taken from the side and top view in basil plants.....	45
2.4.4	Effect of soil Cd concentrations on plant physiological indices generating using HSI taken from the side and top view in kal plants.....	50
2.4.5	Effect of soil Cd concentration on colormaps and reflectance graphs generated using NDVI data collected during the first and third sampling time points.....	55
2.5	Discussion.....	58
2.6	Conclusions.....	63
CHAPTER 3. IDENTIFICATION OF MATHEMATICAL MODELS THAT CAN QUANTIFY CADMIUM CONCENTRATION IN TWO LEAFY GREEN CROPS.....		65
3.1	Abstract.....	65
3.2	Introducion.....	65
3.3	Materials and methods.....	71
3.3.1	Experimental design and plant elemental concentrations at harvest.....	71
3.3.2	Elemental concentrations in aboveground plant biomass.....	71
3.3.3	Manual measurements and images collected during the experiment.....	71
3.3.4	Spectral pre-processing.....	72
3.3.5	Standard normal variation (SNV).....	72
3.3.6	Mathematical models.....	73
3.3.7	Statistical analyses.....	74
3.4	Results and Discussion.....	74
3.4.1	Cadmium and nitrogen concentrations in basil and kale biomass at harvest quantified with traditional chemical extraction and analyses.....	74
3.4.2	Changes in basil and kale leaf area in response to biochar and soil Cd levels.....	76
3.4.3	Spectral pre-processing to develop models to predict Cd concentrations in plant foliage.....	78
3.4.4	Developing mathematical models to estimate Cd concentration in leafy green crops . .....	82
3.5	Conclusions.....	91
CHAPTER 4. CONCLUSION.....		92
BIBLIOGRAPHY.....		95

## LIST OF TABLES

Table 2.1 Fourteen hyperspectral indices with potential to detect cadmium stress in leafy green	31
Table 2.2. Effect of soil cadmium concentrations and plant species on aboveground biomass and elemental concentrations in basil and kale at harvest .....	35
Table 2.3 Effect of soil amendments on aboveground biomass and elemental concentrations at harvest in basil and kale grown in soil subject to one of four soil Cd concentrations.....	36
Table 2.4 Effect of soil amendment and soil cadmium concentration on fourteen hyperspectral indices generated using the side view of images of basil plants collected at three time points....	46
Table 2.5 Effect of biochar and soil cadmium concentration on fourteen hyperspectral indices generated using the top view of images of basil plants collected at three time points. ....	48
Table 2.6 Effect of soil amendment and soil cadmium concentration on fourteen hyperspectral indices generated using the side view of images of kale plants collected at three time point. ....	51
Table 2.7 Effect of soil amendment and soil cadmium concentration on fourteen hyperspectral indices generated using the top view of images of kale plants collected at three time points.....	53
Table 3.1 Statistics table demonstrating the power of these models to detect Cd concentrations in basil and kale aboveground foliage using the transformed data. ....	85

## LIST OF FIGURES

Figure 1.1 Cacao production map in Colombia (adapted from Benjamin et al., 2018).....	17
Figure 2.1 Basil and kale plants grown in soils amended or not with biochar and subject to 0, 5, 10 and 15 ppm soil Cd concentrations. These pictures were taken just prior to harvesting. ....	33
Figure 2.2 Height of basil plants at three time points (a) and dry weight of aboveground biomass after harvest (b) when grown in soil amended with four soil Cd concentrations. Different letters represent significant differences between treatments ( $P<0.05$ ). ....	33
Figure 2.3 Height of kale plants at three time points (a) and dry weight of aboveground biomass after harvest (b) when grown in soil amended with four soil Cd concentrations. Different letters represent significant differences between treatments ( $P<0.05$ ).....	38
Figure 2.4 SPAD readings taken on the leaves of basil plants grown in soil amended with four Cd concentrations at three time points. Different letters represent significant differences between treatments ( $P<0.05$ ).....	40
Figure 2.5 SPAD readings taken on the leaves of kale plants grown in soil amended with four Cd concentrations at three time points. Different letters represent significant differences between treatments ( $P<0.05$ ).....	41
Figure 2.6 Height of basil plants grown with and without biochar (3% v/v) and subject to one of four soil Cd concentrations during three time points. Black bars represent basil plants without biochar (NB) and blue bars represent plants with biochar. Different letters represent significant differences between treatment ( $p<0.05$ ).....	43
Figure 2.7 Height of kale plants grown with and without biochar (3% v/v) at one of four soil Cd concentrations during four time points. Black bars represent basil plants without biochar (NB) and blue bars represent plants with biochar. Different letter represents significant differences between treatment ( $p<0.05$ ).....	44
Figure 2.8 NDVI color maps and reflectance graphs of basil plants grown in soil amended with biochar and subject to 0 or 15 ppm soil Cd during the first and third sampling points. ....	56
Figure 2.9 NDVI color maps and reflectance graphs of kale plants grown in soil amended with biochar and subject to 0 or 15 ppm soil Cd during the first and third sampling point. ....	57
Figure 3.1 Diagram obtained from Shi et al. (2014) illustrating how hyperspectral imaged can be used to conduct experiments to develop models for estimated heavy metal concentrations.....	73
Figure 3.2 Boxplots indicating concentrations of cadmium and nitrogen obtained using ICP-OES in the aboveground biomass of basil and kale plants subject to four soil Cd concentrations. Different letters indicate significant differences at the $P<0.05$ .....	74
Figure 3.3 Figure 3.3 Changes in basil (top panel) and kale (bottom panel) leaf area in plants amended or not with biochar and subject to four soil Cd concentrations at three time points. Different letters indicate significant differences at the $P<0.05$ .....	76

Figure 3.4 Raw reflectance and absorbance data and normalized absorbance generated using a random sample of basil and kale plants subjected to soil Cd stress..... 79

Figure 3.5 Spectral data generated using a random basil or kale plant subject to soil Cd stress. The top panel represents the first derivative and first derivative + SNV absorbance data, and the right panel represents the detrending and detrending + SNV absorbance..... 80

Figure 3.6 ANN regression model generated using the detrending + SNV transformed data of plants subject to soil Cd stress ..... 82

Figure 3.7 Confusion matrices generated using the ANN model and different data transformations using leafy greens subject to Cd stress. .... 83

Figure 3.8 8 PCA and PLS models developed using basil and kale plants subject to soil Cd stress and images collected top view of plants. .... 87

Figure 3.9 PCA and PLS models developed using basil and kale plants subject to soil Cd stress and images collected side view of plants ..... 89

Figure 4.1 NDVI color maps and reflectance graphs of cacao plants grown in soil subject to 0 or 10 ppm soil Cd..... 94

## ABSTRACT

Cadmium (Cd) is a highly mobile and toxic heavy metal that negatively affects plants, soil biota, animals and humans, even in very low concentrations. Currently, Cd contamination of cocoa produced in Latin American countries is a significant problem, as concentrations can exceed acceptable levels set by the European Union (0.5 mg/kg), sometimes by more than 10 times allowable levels. In South America, *Theobroma cacao* is an essential component of the basic market basket. This crop contributes to the Latin-American trade balance, as these countries export cacao and chocolate-based products to major consumer countries such as the United States and Europe. Some soil amendments can alter the bioavailability and uptake of Cd into edible plant tissues, though cacao plants can accumulate Cd without displaying any visible symptoms of phytotoxicity, which makes it difficult to determine if potential remediation strategies are successful. Currently, the only effective way to quantify Cd accumulation in plant tissues is via destructive post-harvest practices that are time-consuming and expensive. New hyperspectral imaging (HSI) technologies developed for use in high-throughput plant phenotyping are powerful tools for monitoring environmental stress and predicting the nutritional status in plants. Consequently, the experiments described in this thesis were conducted to determine if HSI technologies could be adapted for monitoring plant stress caused by Cd, and estimating its concentration in vegetative plant tissues. Two leafy green crops were used in these experiments, basil (*Ocimum basilicum* L. var. Genovese) and kale (*Brassica oleracea* L. var. Lacinato), because they are fast growing, and therefore, could serve as indicator crops on cacao farms. In addition, we expected these two leafy green crops would differ in their morphological responses to Cd stress. Specifically, we predicted that stress responses would be visible in basil, but not kale, which is known to be a hyperaccumulator. The plants were subject to four levels of soil Cd (0, 5, 10 and 15 ppm), and half of the pots were amended with biochar at a rate of 3% (v/v), as this amendment has previously been demonstrated to improve plant health and reduce Cd uptake. The experiments were conducted at Purdue's new Controlled Environment Phenotyping Center (CEPF). The plants were imaged weekly and manual measurements of plant growth and development were collected at the same times, and concentrations of Cd as well as many other elements were determined after harvest. Fourteen vegetation indices generated using HSI images collected from the side and top view of plants were evaluated for their potential to identify subtle signs of plant stress due soil Cd

and the biochar amendment. In addition, three mathematical models were evaluated for their potential to estimate Cd concentrations in the plant biomass and determine if they exceed safe standards (0.28 mg/kg) set by the Food and Agriculture Organization (FAO) for leafy greens. Results of these studies confirm that like many plants, these leafy green crops can accumulate Cd levels that are well above safety thresholds for human health, but exhibit few visible symptoms of stress. The normalized difference vegetation index (NDVI) and the chlorophyll index at the red edge (CI\_RE) were the best indices for detecting Cd stress in these crops, and the plant senescence and reflectance index (PSRI) and anthocyanin reflectance index (ARI) were the best at detecting subtle changes in plant physiology due to the biochar amendment. The heavy metal stress index (HMSSI), developed exclusively for detecting heavy metal stress, was only able to detect Cd stress in basil when images were taken from the top view. Results of the mathematical models indicated that principal components analysis (PCA) and partial least squares (PLS) models overfit despite efforts to transform the data, indicating that they are not capable of predicting Cd concentrations in these crops at these levels. However, the artificial neural networks (ANN) was able to predict whether leafy greens had levels of Cd that were above or below critical thresholds suggested by the FAO, indicating that HSI could be further developed to predict Cd concentrations in plant tissues. Further research conducted in the field and in the presence of other environmental stress factors are needed to confirm the utility of these tools, and determine whether they can be adapted to monitor Cd uptake in cacao plants.

## CHAPTER 1. INTRODUCTION

*Theobroma cacao* is a tropical evergreen tree that has been cultivated throughout history for its great-tasting and highly nutritional seeds. Its seeds are generally processed to produce cocoa powder, cocoa butter and chocolate. An average of 4.81 millions tons of cacao beans are produced worldwide each year (ICCO, 2020). Cocoa is grown commercially in Central America, South America and the Caribbean, as well as in West Africa and tropical Asia. Africa is the largest producer, with 73.1% of world cocoa production, followed by Latin America with 16.9%, and Asia and Oceania with 9.9% (ICCO, 2018) The most notable cacao producing countries in Africa include the Ivory Coast, Ghana, Nigeria, Cameroon, and in Asia, Indonesia and Malaysia. In the Americas, the main cacao producing countries are Brazil, Ecuador and Venezuela, and to a lesser extent, Colombia, Costa Rica, Cuba, Mexico, Peru and the Dominican Republic. However, the type of cacao grown in different region of the world can vary. For example, most of the finer, aroma cacao production occurs in the America's producing 70% of the world's fine or flavor cocoa (Caligiani et al., 2016; FAO, 2013). According to ICCO (International Cocoa Organization) estimates that 95% of cacao from Colombia is exported as fine and flavor. It is found through the Andes toward the low-lands of Venezuela, Colombia, and Ecuador and northward to Central America and Mexico, and to a large number of Caribbean Islands. (Caligiani et al., 2016). Currently, Venezuela, Ecuador, Mexico and Colombia produce the finest aroma cocoa (creole variety) , representing approximately 5% of the total world cacao production (Caligiani et al., 2016). However, creole is difficult to cultivate because it is highly susceptible to diseases.

There are many opportunities to increase production and exportation of this fine aroma cacao to major importing countries like Europe (40%) and the United States (Benjamin et al., 2018). However, the cocoa sector is unstable, and prices can drop quickly due to concentration of the cocoa business in the hands of a few transnational corporations that dominate the trade in raw cacao materials and the confectionery industry (ICCO, 2009). In addition, in many Latin American countries, cacao can be contaminated by cadmium, a toxic heavy metal, which threatens existing production and could limit opportunities for further expansion (Zug et al., 2019). The broad goal of the research described in this thesis, was to identify new technologies that could be used to help address this issue, opening up opportunities for further development of the cacao industry in Colombia.

## 1.1 Cacao history and production

*Theobroma cacao* is a member of the Malvaceae family, but it is also sometimes classified within the Sterculiaceae family (Aprotosoie et al., 2016; Caligiani et al., 2016; de Souza et al., 2018). The name comes from the Greek "Theos" which means god and "broma" which means food. The full name means "the food of Gods". The *Theobroma* genus contains roughly 22 species of small understory trees that are native to the tropical forests of Central and South America (Whitlock and Baum, 1999) These trees are characterized by large, alternate unlobed leaves and small flowers that grow in the leaf axils or directly on the trunk. Flies and midges are the main pollinators, though mosquitoes can also feed on their pollen (Russel et al., 2008). Natural propagation is by seed, and cultivation is possible only in tropical areas. The seeds are roasted and pulverized to make cocoa after the fat has been removed, though if the goal is to make chocolate, then the fat is retained (Caligiani et al., 2016). The fruit is harvested twice a year, in February-March and April-July. However the summer harvest usually produces fruit of better quality (Caligiani et al., 2016) While cacao has been consumed in Latin America for over 1,000 years, chocolate became cherished worldwide around 1850 when food entrepreneurs discovered that by using more cacao butter, it was possible to create a solid form of this delicious bean (World Cocoa Foundation, 2015).

There is a lot of genetic variation among *Theobroma cacao* varieties and individual clonal lines, each producing cacao seed with different flavors, aromatics and bitterness (Aprotosoie et al., 2016) The three major types of cocoa are Forastero, Criollo and Trinitario (Benjamin et al., 2018). Forastero is the most common source of 'bulk' cocoa in the global market (~80% of the world's total cacao market) (Benjamin et al., 2018). Criollo produces the highest quality beans and best chocolate, because its seeds are less bitter and more aromatic than the others (Benjamin et al., 2018)). The beans produced from Criollo yield fine cocoa with greater smoothness and a high lipid content, and white chocolate is made from seeds produced by these trees. Criollo represents less than 5% of the total amount of cocoa produced, because it has lower yields and is susceptible to fungal pathogens and other pests (Benjamin et al., 2018). Trinitario is hybrid of Criollo and Forastero, and it yields some of the best traits from both of these types (Benjamin et al., 2018). Trinitario was developed by The International Cocoa Organization (ICCO). Forastero is commonly found in Africa, while Criollo and Trinitario is common in Latin America, the Caribbean, and some countries in the Atlantic, Indian, and Pacific oceans (Benjamin et al., 2018).

Altitude and temperature are critical considerations for cacao production. Cocoa can only be cultivated in a narrow strip of land running from sea level up to 1,300 meters above sea level that has an average temperature of between 22 and 30 °C and a minimum temperature of 18°C (ICCO, 2013). Cacao trees also require a lot of water and are commonly grown in areas where average rainfall is 2,500 millimeters per year that is evenly distributed throughout the year, and a relative humidity 80% (ICCO, 2013). The trees cannot be exposed to high winds, and windbreaks are highly recommended to prevent damage in areas where this could be an issue (Ruf and zadi, n.d.). Cacao trees thrive in the shade of other more robust trees, and shade greater than 70% is critical during establishment (Ngala, 2015). After the third year, shade requirements can be reduced to 30%. Both excessive shade and shade deficits are detrimental to cacao production, because they can favor the incidence and severity of diseases and pests (Moreno and Sánchez, 1990). Soil health is another critical factor in cacao production. The pH of the soil must be between 5.0 and 7.0 in the surface layer to a depth of one meter, and the soil must have sufficient organic matter to retain nutrients and decrease compaction (Ngala, 2015). There has been a significant amount of research conducted to optimize fertilization, irrigation, draining and planting in cacao production, though in many countries, farmers lack technological training associated with best management practices, which limits productivity (Bot et al., 2005).

## **1.2 Cacao production in Colombia**

In South America, cacao is an essential component of many local economies. The crop contributes to the Latin-American trade balance, as these countries export cacao and chocolate-based products to major consumer countries such as the United States and Europe (Benjamin et al., 2018; ICCO, 2009). In Colombia, cacao has long been an important crop, though with the signing of the recent peace treaty in 2016 that ended a civil war that lasted over 60 years, there are many opportunities to scale up production of this valuable crop. In particular, there is a lot of interest in replacing the cultivation of illicit crops, which flourished during the civil war period, with cacao. In 2015, the U.S. Department of Agriculture (USDA) and the U.S. Agency for International Development (USAID) initiated a project called Cacao for Peace (<https://www.fas.usda.gov/data/colombia-cacao-peace-overview>) to aid in this effort. The Cacao for Peace Project aims to "strengthen key Colombian agricultural institutions for cocoa, in the public and private sectors, with cooperative research, technical assistance, outreach, and

education. The Cacao for Peace vision is to “improve rural well-being through development that is inclusive and sustainable with a positive impact on cacao farmers' incomes, economic opportunities, stability and peace.” (Colombia Purdue partnership)

Efforts like the Cacao for Peace project have appeared to be successful, as acreage devoted to this high value crop have increased over the past ten years. For example, according to the Colombian National Administrative Department of Statistics, DANE ( <https://www.dane.gov.co> ) Production went from 106 thousand hectares planted to 162 thousand hectares in 2015. During 2015, cacao cultivation increased by another 32.6% and in 2017 Colombia produced a total of 60,000 tons of cacao beans (Benjamin et al., 2018). In 2015, exports of cocoa and confectionery products reached US\$333.2 million, which represents 10 % of the total exports of the Colombian agribusiness sector (DANE, 2015). According to the Colombian Ministry of Agriculture, “we are currently in the right moment for the cocoa sector, as planting this crop has become an important post-conflict strategy and there are many opportunities to support the growth of this valuable crop”. Cacao is currently grown in 30 departments of the country. It is an essential axis of the economy in more than 760 municipalities with about 38 thousand producers producing 60500 tons (FEDECACAO, 2018). According to Fedecacao (<https://www.fedecacao.com.co/portal/index.php/es/>), an organization established in 1960 to provide support for small and medium sized cacao farmers in Colombia, the main cocoa-producing department in Colombia is in Santander, which produces 23.3% of the country’s cacao crop (Figure 1.1). The second largest producing department is Arauca with 11.7% of the country’s crop in municipalities such as Arauquita, Saravena, Fortul and Tame. The third most significant producing department is Antioquia with 10% of the crop in the Urabá area. Other departments with smaller but growing cocoa production sectors include Norte de Santander, Nariño, Tolima, Meta, Huila, Boyacá and Cesar.

Since the signing of the peace treaty and efforts to facilitate post-conflict development programs, farmers have returned to their plantations and have embraced the technology taught by the National Cocoa Federation in terms of disease management, which has been a major problem in Colombia. However, awareness of potential problems associated with cadmium (Cd) contamination in many Latin American countries have now become apparent. The German Confederation of Confectioners was one of the first groups to identify the problem. They contacted the embassy of Colombia to report the existence of cadmium in imports from Latin America. New

standards for the maximum accepted level for cadmium (Cd) in beans in cacao beans that was recently established in Europe and is being adopted in many other countries worldwide, is 0.5 mg kg<sup>-1</sup> (Arévalo-Gardini et al., 2017; Cacao for peace, personal communication). However, according to Casa Luker ([www.casaluker.com](http://www.casaluker.com)), a major chocolate company in Colombia, Cd levels in Colombian cacao beans can reach as high as 30 ppm mg kg<sup>-1</sup> (Casa Luker, personal communication).

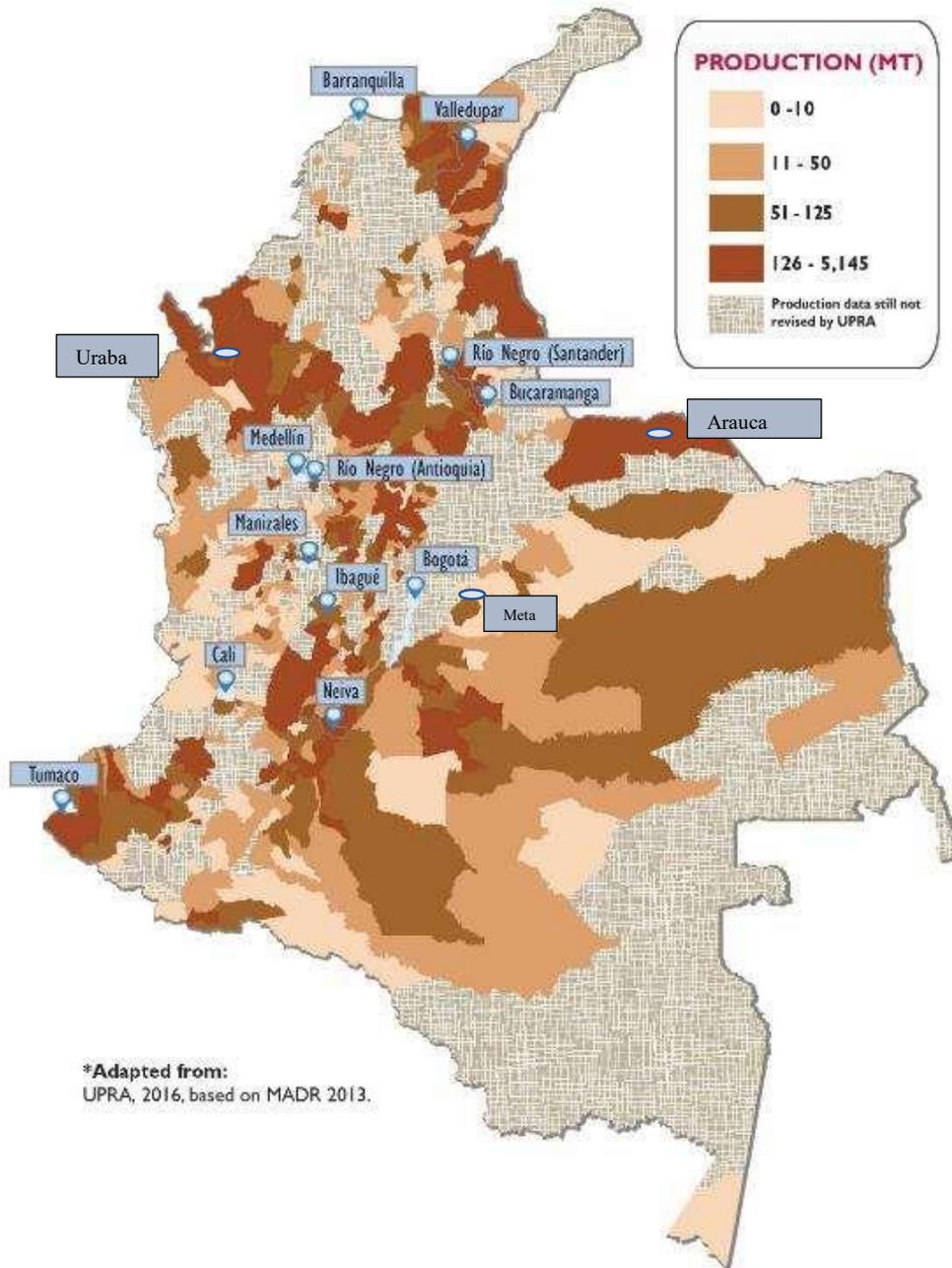


Figure 1.1 Cacao production map in Colombia (adapted from Benjamin et al., 2018)

### 1.3 Cadmium in soils

Cadmium is a heavy metal that is commonly found in the earth's crust in low concentrations. It is ranked seventh in the top-ten priority list of hazardous substances by the US Department of Health and Human Services' Agency for Toxic Substances and Disease Registry (He et al. 2015). This is due to the fact that Cd exposure can lead to renal tubular dysfunction, bone damage, and lung diseases, posing health risks for both humans and animals even when present in very low concentrations (Ismael et al., 2019). The presence of high levels of Cd in soil can also negatively affect critical soil processes like nutrient cycling (He et al., 2015), as well as critical plant processes such as photosynthesis (He et al., 2015). Nevertheless, this heavy metal can accumulate in edible plant tissues like cacao beans, even though plants show no visible symptoms of toxicity (Sánchez-Pardo et al., 2013), making it difficult to know when this element is a problem.

Cadmium can be found in some soils due to natural geogenic processes, but high levels are generally associated with anthropogenic activities such as mining and contaminated fertilizers (He et al., 2017). Once this element is introduced into soils, it is nearly impossible to remove, therefore if possible, identifying sources and preventing the introduction of this heavy metal into cacao farms is a critical step in mitigating the challenge associated with its accumulation in cacao beans. However, once soils become contaminated, like many of the more traditional cacao growing regions in Colombia, understanding factors that affect the bioavailability of Cd in soil is important for mitigation. Not all forms of Cd that are present in soil are in soluble form and therefore are bioavailable for plant uptake. For example, Cd can be present in soil in water-soluble fractions, in interchangeable fractions on soil particles that are bound by carbonates, Fe-Mn oxides, organic matter and sulfur, or in tightly bound residual materials (He et al., 2015). Water-soluble and exchangeable fractions (the Cd pool on cation exchange sites in soil) are generally the most bioavailable, while the remaining fractions are considered the most inactive and least available to biota (He et al., 2015). Cadmium from anthropogenic processes is usually more mobile than geogenic processes (He et al., 2015).

Many biochemical processes can affect the mobility of Cd in soils, including sorption (specific, nonspecific and ligand mediated) on inorganic and organic colloidal materials, precipitation and complexation with organic and inorganic ligands in the soil solution, and reactions such as chelation-dissociation, mineralization, assimilation and protonization-deprotonation (He et al., 2015; Loganathan et al., 2012)). Unlike some other heavy metals,

approximately 70% of Cd present in soil is bound via adsorption, which can result in this element becoming desorbed and migrating to deeper layers of the soil (He et al., 2015). At high concentrations, the dominant mechanism tends to be precipitation with anions such as sulfide ( $S^{2-}$ ), Hydroxide ( $OH^-$ ), carbonate ( $CO_3^{2-}$ ) and phosphate ( $PO_4^{3-}$ ) due to industry (Loganathan et al., 2012; He et al., 2015). When Cd is bound due to a nonspecific sorption, it is usually retained on negatively charged sites on the surfaces of soil colloids by electrostatic attraction, which can undergo ion exchange with other cations in the soil solution (Loganathan et al., 2012). In contrast, specifically sorbed Cd is retained on soil particles due to chemical bonds with neutral and negatively charged sites, and therefore, this fraction is not easily exchanged for other cations. This type of sorption can decrease the net negative surface charge of soil particles, and increase the ground zero net charge point (He et al. 2015). When Cd is in soil solution, factors such as the type and amount of clay and organic matter present, soil pH, presence of other cations and anions, and the percentage of sorption sites occupied by Cd can all influence its bioavailability to plants (He et al., 2015; Araujo et al., 2016). For example,  $Cd^{2+}$  on cation exchange sites (negatively charged exchange sites in soil) such as those on the edges and interlayers of 1:1 and 2:1 layer silicate minerals, respectively (Loganathan et al., 2012) as well as negatively charged carboxylic and phenolic groups in soil organic matter (He et al 2015).

Given this information, soil scientists are working on ways to stabilize and promote long-term Cd sequestration in soils, to prevent uptake into edible plant tissues like cacao beans. For example, amending soil with certain types of biochar has been shown to reduce the uptake of Cd in soil (Gomez and Hoagland, unpublished; Hayyat et al., 2016; Lehmann and Joseph, n.d.)

#### **1.4 Using leafy greens as indicator crops**

Plant species differ dramatically in their ability to tolerate the presence of Cd and accumulate this toxic element in different plant compartments for reasons that are still not well understood. Plant tissue concentrations of 3-30  $mg\ kg^{-1}$  are generally considered toxic for most plant species (Chen et al., 2011; He et al 2015), though some so-called ‘hyperaccumulator’ species, can contain more than 100  $mg\ Cd\ kg^{-1}$  in plant tissues without displaying any visible signs of toxicity (He et al 2015). Cacao crops appear to be able to tolerate high levels of soil Cd without displaying visible symptoms of toxicity (Chavez et al., 2015). Moreover, this is a perennial crop and thus cacao trees are capable of reallocating elements internally over long time periods, making it difficult to

determine if efforts to stabilize Cd in soil are working. Instead, annual leafy greens crops could be used as an indicator crop because they also take up large volumes of Cd and have a relatively high potential for Cd translocation to above ground tissues (Baldantoni et al., 2015). Moreover, they tend to grow quickly and some species like spinach can display severe symptoms of Cd stress (Gomez and Hoagland, unpublished). Finally, these crops are also an important component of the human diet, and can put human health at risk when grown on contaminated soils (Huang et al., 2017). Therefore, soil remediation strategies are also needed for these crops.

Basil is grown worldwide for direct consumption, as well as a medicinal herb used in some medical treatments, and to produce oil for use in flavoring and perfumes (Kumar et al., 2016). Previous reports indicate that factors such as seed germination and early seedling growth of basil plants are severely affected by Cd, especially when levels reach 15-20 ppm of Cd (Gharebaghi et al., 2017). In the latter study, many aspects of the basil plant foliage were affected by Cd, including the number of leaves, plant height, height of the cotyledons following emergence and root length. Consequently, detecting Cd stress in basil by the naked eye is possible. In contrast, many Brassica species like kale have been classified as Cd hyperaccumulators, which can be useful in phytoremediation strategies that use plants to accumulate Cd in their leaves by harvesting the plant biomass to remove Cd from the system (Haghighi et al., 2016). Kale plants have been shown to be tolerant to soil Cd, displaying no visible symptoms of toxicity, and no significant reduction in plant biomass even at high soil Cd concentrations (Haghighi et al., 2016). Consequently, in most cases, it would be impossible to detect Cd uptake in kale crops visibly, making these crops particularly dangerous to grow on contaminated soils.

### **1.5 Using hyperspectral imaging to detect Cd stress and quantify Cd accumulation in edible plant tissues**

Currently, the most effective way to quantify Cd in plant tissues is via post-harvest by extracting the tissues with a chemical procedure followed by analysis such as inductively coupled plasma mass spectroscopy (ICP-MS). However, this procedure is time consuming and expensive, thus most growers and many researchers in Colombia may not have access to this technique. Alternatively, a noninvasive and faster technique such as hyperspectral imaging (HSI) could be developed as a means to quantify Cd stress and estimate Cd concentrations in leaves. Hyperspectral imaging measures reflectance of plant tissues in the visible light and near-infrared light ranges. It

is widely used in studies to detect environmental stress caused by factors that include high temperatures, drought, salinity, flooding, nutritional deficiencies (Bergsträsser et al., 2015; Neilson et al., 2015). In addition, some have begun evaluating its potential to quantify stress caused by pollutants such as heavy metals, herbicides, detergents (Wang et al., 2018; Zhou et al., 2019). Consequently, the goals of the studies described in this thesis, were to 1) determine whether HSI can be used to detect Cd stress in two distinct leafy green crops, and 2) investigate whether HSI can be used to estimate Cd concentrations in the edible tissues of these crops.

## **CHAPTER 2. USING HYPERSPECTRAL IMAGING TO DETECT CADMIUM STRESS IN KALE AND BASIL AND DETERMINE WHETHER BIOCHAR CAN REDUCE UPTAKE INTO EDIBLE PLANT TISSUE.**

### **2.1 Abstract**

Cadmium (Cd) is a heavy metal found naturally in the earth's crust in low concentrations, but it can accumulate in soil and the edible tissues of crops, which can negatively affect soil, plant and human health. Soil amendments like biochar have potential to reduce the bioavailability of Cd in soil thereby reducing uptake in crops and preventing human health risks. However, currently the only effective way to quantify Cd accumulation is via destructive post-harvest practices that are time consuming and expensive. Consequently, the primary goal of this study was to determine whether hyperspectral imaging (HSI) can be used as a non-destructive method to quantify Cd stress and estimate uptake in two distinct leafy green crops (basil and kale) during crop production. In addition, we aimed to characterize how Cd affects the development and health of these crops, and determine whether a locally sourced biochar amendment derived from hardwoods can enhance the health of these plants and reduce Cd uptake into the edible tissues. Results confirm that while these crops can both accumulate levels of Cd that are well above safe thresholds for human health, they generally show few visible symptoms of plant stress. The biochar amendment and rate evaluated in this trial did appear to have some subtle effects on reducing plant stress responses due to Cd in basil, but was not effective in preventing Cd uptake in either crop species. Several vegetative indices including the normalized difference vegetation index (NDVI) and chlorophyll index at red edge (CI\_RE) appear to have the potential to reveal Cd stress in these crops. Also, the plant senescence and reflectance index (PSRI) and the anthocyanin reflectance index (ARI) appear to have potential to detect subtle changes in plant physiology due to Cd and biochar amendments. However, further research conducted in the field and in the presence of other environmental stress factors are needed to confirm the utility of these tools.

### **2.2 Introduction**

Cadmium (Cd) is a heavy metal found in the earth's crust and in some soils used to grow crops. This metal is ranked seventh among the top ten-priority list of hazardous substances by the

US Department of Health and Human Services' Agency for Toxic Substance and Disease Registry (He et al., 2015) because it can cause many health problems in humans, even in very small concentrations. For example, Cd can cause renal tubular dysfunction, bone damage and lung diseases in humans (Ismael et al., 2019). One of the biggest sources of Cd for humans is plant-based foods. It has been estimated that plant consumption can contribute from 70 to more than 90% of the total human intake of Cd (Baldantoni et al., 2015). The biggest challenge in preventing this human health risk, is that plants can accumulate concentrations of Cd that are toxic to humans, while appearing to be healthy and showing no toxicity symptoms (Ismael et al., 2019; Sánchez-Pardo et al., 2013). Thus, farmers may continue to grow, harvest and sell contaminated produce without being aware that they are putting their customers at risk.

The presence of Cd in soil can come from both geogenic and anthropogenic sources. Geogenic sources include Cd release following weathering of sedimentary rocks. For example, black shales may contain more than 200 mg kg<sup>-1</sup> Cd, and phosphate rocks can contain approximately 25 mg kg<sup>-1</sup> ( He et al., 2015). Anthropogenic sources include mine waste and industrial processes such as battery production and disposal (He et al., 2015). Agricultural inputs can also be a source of Cd. For example, in South America, contaminated phosphate fertilizers have been identified as the principal source of Cd contamination in cacao production systems (Daniel Bravo, personal communication). Soils with concentrations above 1 mg kg<sup>-1</sup> of Cd indicate anthropogenic sources, since levels this high due to geogenic sources are rare (He et al., 2015). Anthropogenic sources are estimated to contribute ten times more Cd than geogenic sources (He et al., 2015). Therefore, identifying and preventing anthropogenic sources of Cd into agricultural systems represents the best way to protect human health.

Once soils become contaminated with heavy metals like Cd, remediation options are limited because heavy metals will not degrade or decompose over time. Consequently, scientists are investigating factors that can immobilize Cd in soil, thereby preventing uptake into plants. Heavy metals in soils and sediments can be divided into various binding phases, which affect their bioavailability for plants. Metals may be bound to the surface of soil particles in several ways, complexed with ligands in solution, or be present as free ions in solution. Many physicochemical and mineralogical factors can affect the binding activity of heavy metals in soils. Soil pH in particular, is well known to be a critical factor affecting the speciation, solubility and bioavailability of many heavy metals in soil. Cadmium solubility in soil tends to exhibit a sigmoid-

like behavior within a range of pH 4 and 8 (He et al., 2015). Cd is mobile in the pH range of 4-6 in soil, while at higher pH values it converts to insoluble carbonate and phosphate forms. An indirect linear relationship between soil pH and bioavailability or plant uptake of Cd exists; as pH decreases, Cd uptake by plant increases (He et al., 2015)

Soil organic matter and the amount and type of clay particles present are also important factors affecting the bioavailability of Cd due to the potential for adsorption on cation exchange sites present on the surface of these particles (Olaniran et al., 2013). As a result, organic soils tend to have a sorption affinity for Cd that is up to 30 times greater than mineral soils (He et al. 2015). In contrast, soils in tropical regions tend to be acidic and low in organic matter, and their mineralogy is dominated by kaolinite (1:1 clay with only cation exchange sites on the clay edges), which makes them have a low cation exchange capacity (CEC) (Fontes et al., 2006; Hayyat et al., 2016). Consequently, soils like these when contaminated with metals can be phytotoxic to plants because of their inability to retain heavy metals that exist as cations like Cd (Melo et al., 2011). Finally, the presence of salts can alter Cd bioavailability by producing complexes such as  $CdCl_n^{2-n}$  (Filipovic et al., 2018), as can the composition of soil microbial communities, because microbes regulate speciation and solubility of Cd via several processes including redox reactions (Ma et al., 2016).

Because soil amendments can alter soil physical, chemical and biological properties, they have potential to alter Cd bioavailability, and therefore, uptake into edible plant tissues. For example, amending soils with lime can increase soil pH, thereby reducing bioavailability of Cd in soils (Ramtahal et al., 2018). Some studies have also demonstrated that amending soils with composts that are high in organic matter can also reduce Cd uptake into plants such as leafy greens (Gomez and Hoagland, unpublished). However, the organic materials in composts can decompose, and therefore the long-term success of this remediation strategy is unclear. Consequently, another remediation option being investigated is amending soils with biochar, which is defined as a carbon-rich product obtained when biomass such as wood, manure or leaves is heated at temperatures between 350-450 C in the absence of air (Lehmann and Joseph, 2009). Biochar is extremely stable and can last for centuries in soil (Lehmann and Joseph, 2009). Amending soils with biochar can improve soil and plant health (Shoaf et al., 2016), and some biochars have the additional benefit of making heavy metals less mobile via several physicochemical and biological processes (Joseph and Lehman, 2015). For example, when organic materials are subject to

pyrolysis to produce biochar, the surface area and number of cation exchange sites on the surface of the organic materials increases (Joseph and Lehman, 2015). Consequently, amending soils with biochar can cause an increase in pH due to the presence of additional ion exchange sites on biochar surfaces, as well as electrostatic interactions between positively charged heavy metals (Hayyat et al., 2016). This allows biochar to act as a sorbent for solution-phase metals. However, the feedstock and pyrolysis temperatures used to make biochar dramatically alters its physical and chemical properties, and therefore, the potential to alter soil properties and plant health (Shoaf et al., 2016), as well as bioavailability of heavy metals in soil (Lehmann and Joseph, n.d.). Biochar produced from hardwoods tend to have greater surface area and CEC sites than those produced from materials such as grass or manure (Lehmann and Joseph, 2015.). Biochar produced from hardwoods have been shown to increase the success of ryegrass germination in soils with Cd contamination (Beesley et al. 2010), thus have potential for remediation of Cd contaminated soils.

Other factors that can affect soil bioavailability and uptake of Cd into edible plant tissues include plant species and genotype (Ismael et al., 2019). In many plants, concentrations of heavy metals are generally higher in roots than shoots because of negative ion exchanges sites on roots (Baldantoni et al., 2015). However, leafy vegetables have a relatively high potential for Cd absorption and translocation into aboveground tissues, thus are considered Cd accumulators (Baldantoni et al. 2015). As leafy vegetables are an important component of the human diet, they can be a significant risk factor in areas with high soil Cd concentrations, so identifying effective remediation strategies is critical. However, scientists need effective tools to determine if remediation strategies are effective. Currently, the most effective and common way to quantify the presence of heavy metals in plants tissues is through the use of wet post-harvest chemical methods such as ICP-OES, but these techniques are laborious and expensive (Zhou et al., 2018). Consequently, identifying alternative approaches to quantify Cd uptake in plants would be valuable for scientists and farmers.

Some plants, including leafy greens like spinach and basil, are highly sensitive to the presence of Cd and show clear morphological responses such as stunting, chlorosis, necrosis, blackening of the roots and even death as soil concentrations increase (Gomez and Hoagland, unpublished). This is because Cd can interfere with important physiological processes such as photosynthesis and respiration, and absorption, transport and assimilation of mineral nutrients and water (He et al., 2015). In addition, Cd can interfere with gene and protein expression, inducing or

inhibiting the activity of enzymes, increasing the accumulation of reactive oxygen species, causing lipid peroxidation and altering plant metabolism (He et al., 2017). For example, in *Abelmoschus esculentus* L. and *Zea mays*, Cd has been demonstrated to decrease transcription of photosynthesis-related genes and inactivate enzymes involved in CO<sub>2</sub> fixation and chlorophyll biosynthesis (He et al., 2017; Sharma et al., 2010). Cadmium can also affect nitrogen (N) metabolism by inhibiting nitrate absorption and reducing the activity of enzymes involved in nitrate assimilation pathways (He et al. 2017). However, in some plant species, particularly Brassicas, plants may not exhibit any negative effects of Cd even though they may be experiencing some of these physiological stress responses. For example, kale is an important leafy green in the human diet that seems to be tolerant of Cd, displaying no visibly symptoms of toxicity or significant reductions in plant biomass when grown in contaminated soils. Thus, scientists and farmers cannot always rely on morphological symptoms alone to identify “hot spots” of contamination, or determine if remediation strategies are working.

Imaging technologies that use various wavelengths to detect differences in key plant physiological processes represent one approach that could be used to help researchers and farmers quantify Cd without relying on destructive sampling. For example, SPAD meters are commonly used to quantify chlorophyll concentrations and estimate foliage N based on the close relationship between chlorophyll and foliage N concentrations (Xiong et al., 2015). Since heavy metals like Cd can interfere with photosynthesis and nitrogen uptake and metabolism, SPAD meters could be helpful in identifying Cd stress. However, a nonlinear relationship between chlorophyll and foliage N at high N levels has been shown (Huang et al., 2014), indicating that SPAD meters are not always reliable in different situations. Instead, other imaging technologies that separate light into individual wavelengths could provide better insights. For example, machine vision systems based on this technology are being used to quantify plant populations and plant physiology and stress responses (Lowe et al., 2017). These systems rely on hyperspectral images that can be used to non-destructively assess plant growth rates or morphological changes. These employ digital cameras with subsequent software image analysis that allows for faster and more accurate determination of measurements. Hyperspectral imaging (HSI) in particular, detects hundreds of contiguous narrow wavelength bands in a broad spectral range (Lowe et al., 2017). This technique analyzes a wide spectrum of light instead of just assigning primary colors (red, green, blue) to each pixel. It works in the visible (VIS), NIR, and SWIR regions. VIS-NIR stands for the visible near-infrared regions

that go from 400 nm to 1400 nm, and SWIR stands for the short-wavelength infrared region and goes from 1400– 3000 nm (Huber et al., 2014). These wavelengths can capture changes in leaf pigmentation (400–700 nm), mesophyll cell structure (700–1300 nm) and water content of plants (1300–2500 nm).

During HSI, the light striking each pixel is broken down into many different spectral bands in order to provide more information on what is being imaged, a process that produces large amounts of data. Because these large datasets are so complex and difficult to decipher, some researchers instead focus on a small number of wavelengths that can quantify specific changes in plant responses and be used to develop vegetation indices. Many vegetation indices have been established and each uses a different set of wavelengths to quantify physiological attributes of vegetation that can include general properties of the plant or specific parameters of its growth (Lowe et al., 2017). These indices have also been used to detect early stages of stress symptoms due to heavy metals. For example, the normalized difference of vegetation indices (NDIV) is commonly used to estimate chlorophyll content for photosynthetic efficiency (Humplík et al., 2015), and has also been demonstrated to have some value in detecting heavy metal stress (Zhou et al., 2018). While some of these imaging technologies and associated indices are being used in field settings, lighting can alter the quality of the images and plants can be subject to multiple stress factors. Consequently, it can be difficult to identify indices that best measure individual plant stress factors due to all the confounding factors present in the field.

Comprehensive, high-throughput phenotyping facilities can overcome challenges associated with HSI conducted in the field by providing researchers with the opportunity to combine automated, simultaneous and non-destructive analysis of plant growth, morphology and physiology, to develop a more complete picture of plant growth and vigor throughout the plants life cycle (Neilson et al., 2015). These systems also allow researchers to conduct controlled studies that focus on the evaluation of individual plant stress components. Nevertheless, there is still much work to be done to optimize approaches to evaluating HIS in these phenotyping facilities. For example, while many of these facilities use controlled lighting, the images generated can present specific problems because overlapping leaves, twisting and curving can distort determination of the growth area. In addition, distortion can be generated when the hyperspectral image is taken only from one view (for example, from the top view), or the quality of the camera used is

inadequate. A configuration that introduces more projections (i.e., side views in addition to top views) on the phenotyping platforms can partially solve this problem.

The primary goal of this study was to determine whether HSI can be used as a non-destructive method to quantify Cd stress and uptake in leafy greens crops. In addition, we aimed to characterize how Cd affects the development and health of two distinct leafy green crops, and determine whether a locally sourced biochar amendment derived from hardwoods can enhance plant health and reduce Cd uptake into the edible tissues of these crops. We predicted that increasing levels of Cd would cause morphological responses in basil that could be observed by the naked eye, while Cd stress in kale would only be apparent using HSI imaging. We also predicted that the biochar amendment could reduce Cd uptake levels from soil, but overall reduction would become less at higher concentrations. To test these hypotheses, soils were amended with biochar or left untreated, and subject to three levels of Cd that are all above safe limits for human toxicity, and have been observed to produce visual differences in morphology among some, but not all, of the leafy green crops in previous experiments.

## **2.3 Materials and Methods**

### **2.3.1 Experimental design**

The experiment was set up with a total of 64 pots containing a growth media in equals parts by volume (1:1:1 sand: soil: BM8 potting mix) at Purdue University's Controlled Environment Phenotyping Facility (CEPF) in West Lafayette Indiana, U.S. during spring 2019. The water holding capacity of the growth media was determined by adding increasing amounts of water until the media was fully saturated and water started dripping from the pots. This equated to 1200ml of water in a 4.2 ml pot. Then, the pots were weighed dry (3.9 kg) and wet (5 kg) to know when pots needed supplemental water. During the plant growth phase described below, each pot was weighed daily and watered as needed to maintain adequate moisture, with increasing rates of water as the plants grew over time.

In preparation for the experiment, half of the pots were amended with a locally sourced biochar made from a mix of hardwoods at a rate of 3% v/v, and the other half were left untreated. Individual pots were also amended with CdCl<sub>2</sub> to obtain concentrations of total soil Cd of 0, 5, 10, and 15 ppm. For each Cd concentration, CdCl<sub>2</sub> (Sigma Aldrich) was diluted in sterile water to

obtain the amount needed in 1 ml of solution to ensure that the volume of the growth media in the pot (3.9 kg) would be at the appropriate concentration. One ml of this mixture was added to the pots using a pipettman, and was stirred around the pot using a bamboo stake to mix into the soil. Then all of the pots were saturated with water (1200 ml) and allowed to equilibrate for two weeks to facilitate Cd adsorption onto soil particles.

After the incubation period, half of the pots were planted with basil (cv. Genovese basil), and half with kale (cv. *Lacinato kale*), that had been sown four weeks earlier in potting media (Berger, Ca). Each plant species X soil amended with or without biochar X Cd rate was replicated four times. The pots were set up in a completely randomized design on separate belts for each crop species and sent to the growth chamber where they were subject to 14 hours of daylight and 10 hours of night, 25° C, and 60 % humidity. The plants were watered daily with a mixture of water and 10 ppm Peters 15-5-15 Ca-Mg fertilizer until they were destructively harvested to collect aboveground biomass after approximately 3 months. Specifically, basil plants matured earlier (when they started to flower) and were harvested 62 days after transplanting, and kale plants were harvested 84 days after transplanting. All plant materials were dried in an oven at 60 C for 3 days to obtain dry biomass, and ground to 1 mm size using a UDY cyclone sample mill (UDY Corp., Boulder, Col) for elemental analyses described below.

### **2.3.2 Manual plant measurements**

The height and width of each plant were physically quantified using a ruler at three time points for basil (April 1<sup>st</sup>, April 13<sup>th</sup>, April 18<sup>th</sup>), and four time points for kale (April 2<sup>nd</sup>, April 12<sup>th</sup>, April 25<sup>th</sup>, May 9<sup>th</sup>). The chlorophyll content was also estimated at each of these time points using a SPAD chlorophyll meter (Konica Minolta, inc., New Jersey, U.S.A), which measures the difference between the transmittance of red (650 nm) and infrared (940 nm) light through leaves over a specific leaf area (Uddling et al., 2007; Yuan et al., 2016). During each sampling event, SPAD readings were taken on the third leaf from the top of each plant to ensure that the leaves were approximately the same age during each sampling event.

### **2.3.3 RGB and hyperspectral imaging**

At the same time that the first three manual measurements of each plant species described above were conducted, each plant was also imaged using two cameras. One of these cameras collected red, green and blue (RGB) wavelengths (visible spectrum: 400-700 nm) (Aris, Eindhoven, Netherlands), and the other collected hyperspectral wavelengths. The hyperspectral camera (Middleton Spectral Vision, Middleton, Wisconsin, USA) is a VIS-NIR camera capable of sweeping from 400 to 998 nm wavelengths, and has a spectral resolution of 473 bands. The images from the hyperspectral camera were processed using MATLAB . The orientation and height of each plant was determined in the RGB imaging booth; this information was used in managing the hyperspectral imaging process when each plant entered the hyperspectral imaging booth. Irrigation scheduling and experiment management were carried out with programs developed at Purdue.

### **2.3.4 Spectral processing and vegetative indices**

To increase the quality of the hyperspectral images, the images were pre-processed to remove environmental and physiological factors (noise). Consequently, spectral bands representing images in the 500 nm to 980 nm range were retained, and bands representing the 400 nm to 500 nm range and above 980 nm range were discarded. In addition, the images were subject to white and black referencing to calculate the target reflectance. The hyperspectral images were then used to create values for fourteen different indices that have previously been developed to quantify differences in plant physiological parameters (Table 2.1).

Table 2.1 Fourteen hyperspectral indices with potential to detect cadmium stress in leafy green

Vegetation index	Formulas	Description	Application	Citation
NDVI	$\frac{R_{800} - R_{680}}{R_{800} + R_{680}}$	Normalized difference vegetation	Green Biomass, leaf area index (LAI)	(Sims and Gamon, 2002; Yu et al., 2018)
NBNDVI	$\frac{R_{850} - R_{680}}{R_{850} + R_{680}}$	Narrow-band normalized difference vegetation index	Green Biomass	Sims and Gamon, 2002; Yu et al., 2018)
PRI	$\frac{R_{570} - R_{530}}{R_{570} + R_{530}}$	Photochemical reflectance pigment	Physiology, photosynthesis	(Peñuelas et al., 1997)
NRI	$\frac{R_{570} - R_{670}}{R_{570} + R_{670}}$	Nitrogen reflectance index	Nitrogen	(Huang et al., 2014)
TCARI	$3[(R_{700}-R_{670})-0.2(R_{700}-R_{550})R_{700}/R_{670}]$	Transformed chlorophyll absorption and reflectance index	Chlorophyll, LAI	(P. Lin et al., 2012)
SIPI	$\frac{R_{800} - R_{445}}{R_{800} + R_{445}}$	Structure insensitive pigment index	Pigment ratio between carotenoid and chlorophyll a. Canopy stress and LAI	Yu et al., 2018
PSRI	$\frac{R_{680} - R_{500}}{R_{750}}$	Plant senescence / reflectance index	Senescence	7/30/2020 1:30:00 PM
PhRI	$\frac{R_{550} - R_{531}}{R_{550} + R_{531}}$	Physiological reflectance index	Solar utilization efficiency during development. Discriminate disease and abiotic stress	(Huang et al., 2018)
NPCI	$\frac{R_{680} - R_{430}}{R_{680} + R_{430}}$	normalized pigment chlorophyll index	Chlorophyll	Huang et al., 2014
ARI	$\frac{R_{550} - 1}{R_{700} - 1}$	Anthocyanin reflectance Index	Anthocyanin	(Gitelson et al., 2003)
NDVI_RE	$\frac{R_{705} - R_{705}}{R_{705} + R_{705}}$	Normalized difference vegetation at the red edge	Chlorophyll	(Sun et al., 2019)
CI_RE	$(R_{800} - R_{705}) - 1$	Chlorophyll index at red edge	Chlorophyll	Zhang et al., 2018
MSR_RE	$\left[ \frac{\left( \frac{R_{800}}{R_{670}} \right) - 1}{\left( \frac{R_{800}}{R_{670}} \right) + 1} \right]^{1/2}$	Linearize the relationship between the index and biophysical parameters	Chlorosis	(Ashourloo et al., 2014)
HMSSI	$\frac{CI_{red} - edge}{PSRI}$	Difference in heavy metal stress	Heavy metal	Zhang et al., 2018

### **2.3.5 Colormaps and reflectance graphs**

To evaluate potential correlations between plant health and soil Cd concentrations, and estimate where plants may be accumulating Cd (i.e., old vs. young leaves), colormaps and reflectance graphs were developed using NDVI reflectance data collected from plants that had received biochar and were subject to 0 and 15 ppm soil Cd concentrations during the first and last sampling time points. Each image represents one random plant from each plant species x soil Cd concentration.

### **2.3.6 Elemental concentrations in aboveground plant biomass**

Total carbon (C) and nitrogen (N) in kale and basil aboveground biomass was quantified after subjecting 0.5 g samples of dry biomass to combustion at 840 C (LECO, CE Elantech, Lakewood, NJ, USA). Concentrations of total Cd as well as aluminum (Al), arsenic (As), barium (Ba), beryllium (Be), bismuth (Bi), boron (B), calcium (Ca), cadmium (Cd), caesium (Cs), chromium (Cr), cobalt (Co), copper (Cu), gallium (Ga), indium (In), iron (Fe), lead (Pb), lithium (Li), magnesium (Mg), manganese (Mn), nickel (Ni), phosphorous (P), potassium (K), rubidium (Rb), selenium (Se), silicon (Si), silver (Ag), sodium (Na), strontium (Sr), sulfur (S), tellurium (Te), thallium (Tl), vanadium (V) and zinc (Zn) in plant tissues were determined using ICP-OES (Shimadzu ICPE-9820 and location) following digestion using a Mars 6 (CEM, Charlotte NC, USA) with Xpress vessels. Briefly, 0.5 g samples were placed in 10 ml HNO<sub>3</sub> and subject to a temperature of 200 °C, a pressure of 800 psi, and a power of 900-1050 watts.

### **2.3.7 Statistical analysis**

All statistical analyses (F-tests, ANOVA's and T-tests) were performed using R-studio, MATLAB and Excel 2016 (Microsoft Inc). A deep learning tool was used to create hyperspectral indices within the MATLAB. A general linear model (one-way ANOVA) was used to evaluate individual differences in among the two leafy green species, Cd rate and soil amendment with biochar. Then if the interaction between two factors was not significant, a two-way ANOVA was applied to quantify differences across both of these factors. All multiple comparisons (t-tests) were conducted using Fisher's LDS ( $p < 0.05$ ).

## 2.4 Results



Figure 2.1 Basil and kale plants grown in soils amended or not with biochar and subject to 0, 5, 10 and 15 ppm soil Cd concentrations. These pictures were taken just prior to harvesting.

### 2.4.1 Influence of soil Cd concentration on the development, and dry weight, elemental concentrations and SPAD readings of basil and kale

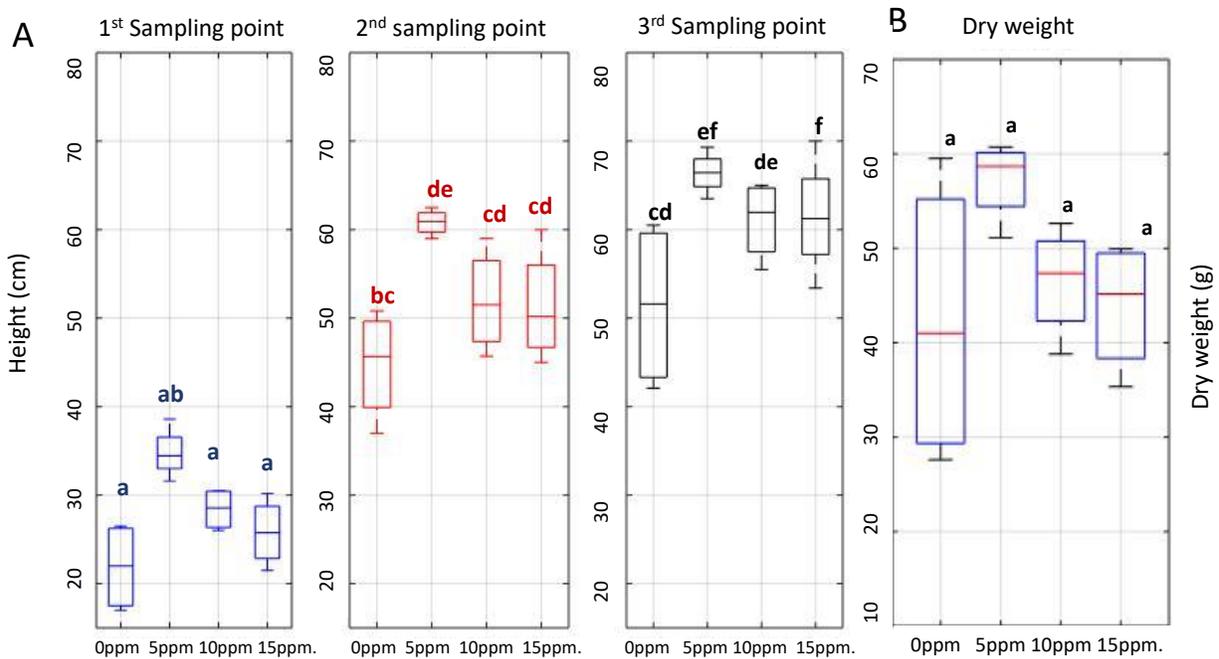


Figure 2.2 Height of basil plants at three time points (a) and dry weight of aboveground biomass after harvest (b) when grown in soil amended with four soil Cd concentrations. Different letters represent significant differences between treatments ( $P < 0.05$ ).

Visible changes in plant development and morphology were observed in basil plants in response to increasing levels of soil Cd (Figure 2.1). In particular, basil plants grown in soil amended with 5 ppm Cd appeared to flower earlier than plants subject to higher soil Cd concentrations (Figure 2.1), and were significantly taller than the control treatment (0 ppm Cd) during the second and third sampling times (Figure 2.2). There were also significant differences in plant height over time among plants subject to individual soil Cd levels (Figure 2.2). Specifically, plants grown in 5, 10 and 15 ppm Cd had significantly greater height between each sampling point, whereas the control plants grown in soil with 0 ppm Cd increased between the first and second time point, but not between the second and third. Basil plants grown in soil with 5 ppm soil Cd also had the highest dry weight of aboveground biomass at harvest, but were not statistically significantly different control plants or those grown in soil with 10 or 15 ppm Cd (Figure 2.2). There were also no differences in the percentage of total C and N or the C:N ratio in basil biomass subject to the different soil Cd treatments, or in Zn concentrations (Table 2.2). In contrast, Cd concentrations were significantly different in the aboveground biomass of all basil plant, and increased concentrations directly corresponded with increasing soil Cd levels (Table 2.3).

Table 2.2. Effect of soil cadmium concentrations and plant species on aboveground biomass and elemental concentrations in basil and kale at harvest

Species	Cd levels	Dry wt. (mg)	Carbon %	Nitrogen %	C:N ratio	Zn (ppm)	Cd (ppm)
Basil	0	43.38	42.10	2.05	21.12	53.00	0.20
Basil	5	51.23	41.18	1.57	30.11	83.70	1.34
Basil	10	47.15	41.97	2.13	20.21	49.93	2.42
Basil	15	43.34	44.68	2.32	20.70	71.06	2.93
Kale	0	64.57	41.47	1.73	24.30	35.31	0.10
Kale	5	66.60	41.49	1.75	24.06	55.65	2.06
Kale	10	72.54	43.53	1.96	22.87	28.45	2.56
Kale	15	75.01	40.69	1.59	25.75	65.28	3.45
Basil		46.28	42.48	2.02 a	23.04	64.42	2.04 b
Kale		69.68	41.80	1.76 b	24.25	46.17	1.72 a
	0	53.98	41.79	1.89	22.71	44.16	0.15d
	5	58.92	41.34	1.66	27.09	69.67	2.05 c
	10	59.85	42.70	2.05	21.45	39.19	2.55 b
	15	59.18	42.68	1.96	23.23	68.17	3.45 a
<i>Statistics</i>	<i>Model</i>	NS	NS	S	S	NS	S
	<i>Species</i>	NS	NS	S	NS	NS	S
	<i>Cd level</i>	NS	NS	NS	S	NS	S
	<i>Spp. X Cd level</i>	NS	NS	NS	S	NS	NS

\*Values with different letters represent significant differences as indicated by LSD ( $p < 0.05$ )

Table 2.3 Effect of soil amendments on aboveground biomass and elemental concentrations at harvest in basil and kale grown in soil subject to one of four soil Cd concentrations

Species	Soil trt	Cd level	Dry wt. (mg)	Carbon	Nitrogen	C:N ratio	Cd	Zn
Basil	Control	0	42.29	41.89	2.11	20.73	0.23	54.83
Basil	Control	5	57.30	39.87	1.48	32.51	1.29	125.10
Basil	Control	10	46.53	41.67	2.13	19.67	2.40	48.75
Basil	Control	15	43.91	42.13	1.96	21.79	3.16	86.33
Basil	Biochar	0	44.48	42.31	1.99	21.52	0.17	51.18
Basil	Biochar	5	45.16	42.50	1.67	27.71	1.39	42.30
Basil	Biochar	10	47.78	42.27	2.12	20.75	2.44	51.10
Basil	Biochar	15	42.76	47.22	2.69	19.62	2.69	55.80
	Control		45.04	43.58	2.12	22.40	1.77	50.09
	Biochar		47.51	41.39	1.92	23.67	1.67	78.75
		0	43.38	42.10	2.05	21.12	0.20 d	53.00
		5	51.23	41.18	1.57	30.11	1.34 c	83.70
		10	47.15	41.97	2.13	20.21	2.42 b	49.93
		15	43.34	44.68	2.32	20.70	2.93 a	71.06
	<i>Statistics</i>	<i>Model</i>	NS	NS	NS	NS	S	NS
		<i>Soil Trt.</i>	NS	NS	NS	NS	NS	NS
		<i>Cd level</i>	NS	NS	NS	NS	S	NS
		<i>Soil Trt. X Cd level</i>	NS	NS	NS	NS	NS	NS
Kale	Control	0	63.62	41.17	1.71	24.71	0.11	41.15
Kale	Control	5	72.95	41.05	1.63	25.47	2.21	76.68
Kale	Control	10	79.99	41.49	1.77	23.50	2.92	35.55
Kale	Control	15	71.81	40.93	1.60	25.60	3.55	44.73
Kale	Biochar	0	65.53	41.77	1.76	23.88	0.08	29.48
Kale	Biochar	5	60.26	41.93	1.86	22.65	1.90	34.63
Kale	Biochar	10	65.09	46.26	2.21	22.04	2.20	21.35
Kale	Biochar	15	78.22	40.45	1.57	25.91	3.34	85.83
	Control		67.27	42.36	1.82	23.73	2.20	42.82
	Biochar		72.09	41.16	1.68	24.82	1.88	49.53
		0	64.57	41.47	1.73	24.30	0.10 c	35.31
		5	66.60	41.49	1.75	24.06	2.06 b	55.65
		10	72.54	43.53	1.96	22.87	2.56 ab	28.45
		15	75.01	40.69	1.59	25.75	3.45 a	65.28
	<i>Statistics</i>	<i>Model</i>	NS	NS	S	S	S	NS
		<i>Soil Trt.</i>	NS	NS	NS	NS	NS	NS
		<i>Cd level</i>	NS	NS	NS	NS	S	NS
		<i>Soil Trt. X Cd level</i>	NS	NS	NS	NS	NS	NS

\*S refers to significant differences (p<0.05)

Stress symptoms in kale were impossible to detect by the naked eye (Figure 2.1, Figure 2.3). In addition, neither aboveground dry biomass (Figure 2.3) or total percent C and N, C:N ratio and Zn in aboveground biomass (Table 2) at harvest were significantly different for plants grown in soil amended with the four soil Cd treatments. Plant height was not significantly different between soil Cd treatments during individual time points, and only the 15-ppm Cd treatment was different between the second and third sampling points (Figure 2.3). Like basil, Cd concentrations in the aboveground biomass of kale increased with increasing soil Cd levels (Table 2.2).

The percentage of N in the aboveground biomass of basil at harvest was significantly higher than kale, and the concentration of Cd was significantly greater in kale than basil (Table 2.3).

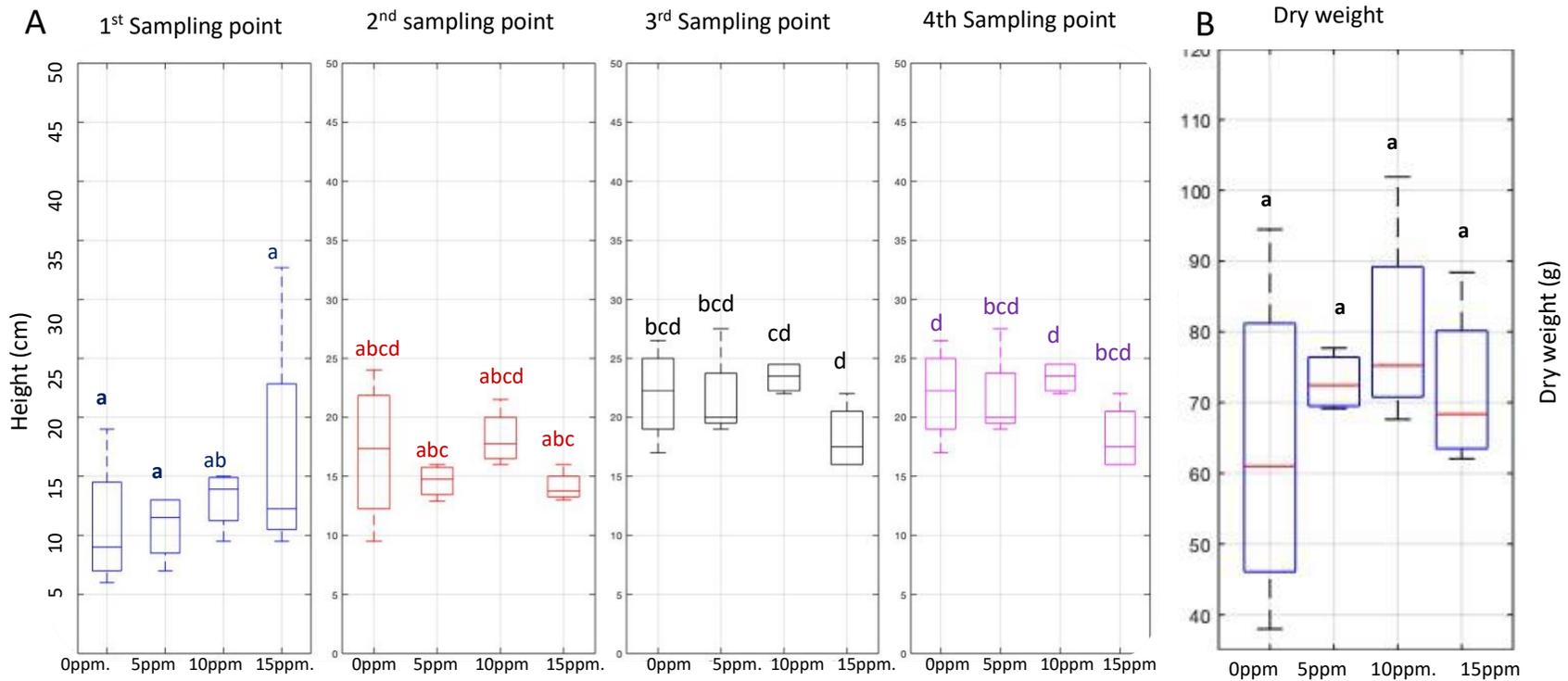


Figure 2.3 Height of kale plants at three time points (a) and dry weight of aboveground biomass after harvest (b) when grown in soil amended with four soil Cd concentrations. Different letters represent significant differences between treatments ( $P < 0.05$ )

Chlorophyll content estimated using a SPAD meter was not significantly different among basil plants subject to individual soil Cd levels, however, there were a few subtle differences over time (Figure 2.4). Specifically, in the absence of biochar, the 5 ppm soil Cd treatment was significantly lower in the third than the second sampling time point, and the 10 ppm soil Cd treatment amended with biochar was also significantly lower in the third relative to the second time point. For kale, there were no significant differences in SPAD readings (Figure 2.5)

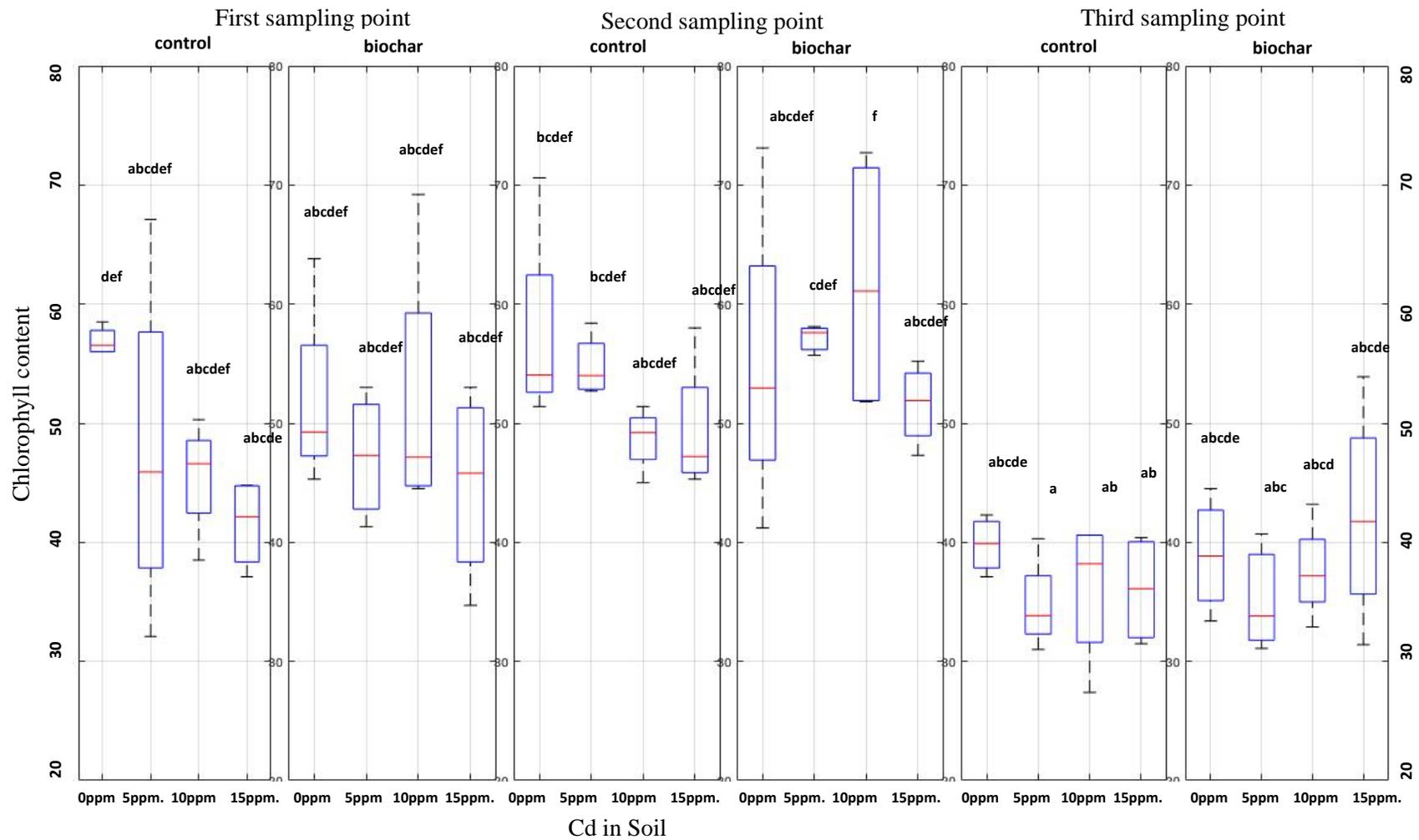


Figure 2.4 SPAD readings taken on the leaves of basil plants grown in soil amended with four Cd concentrations at three time points. Different letters represent significant differences between treatments ( $P < 0.05$ )

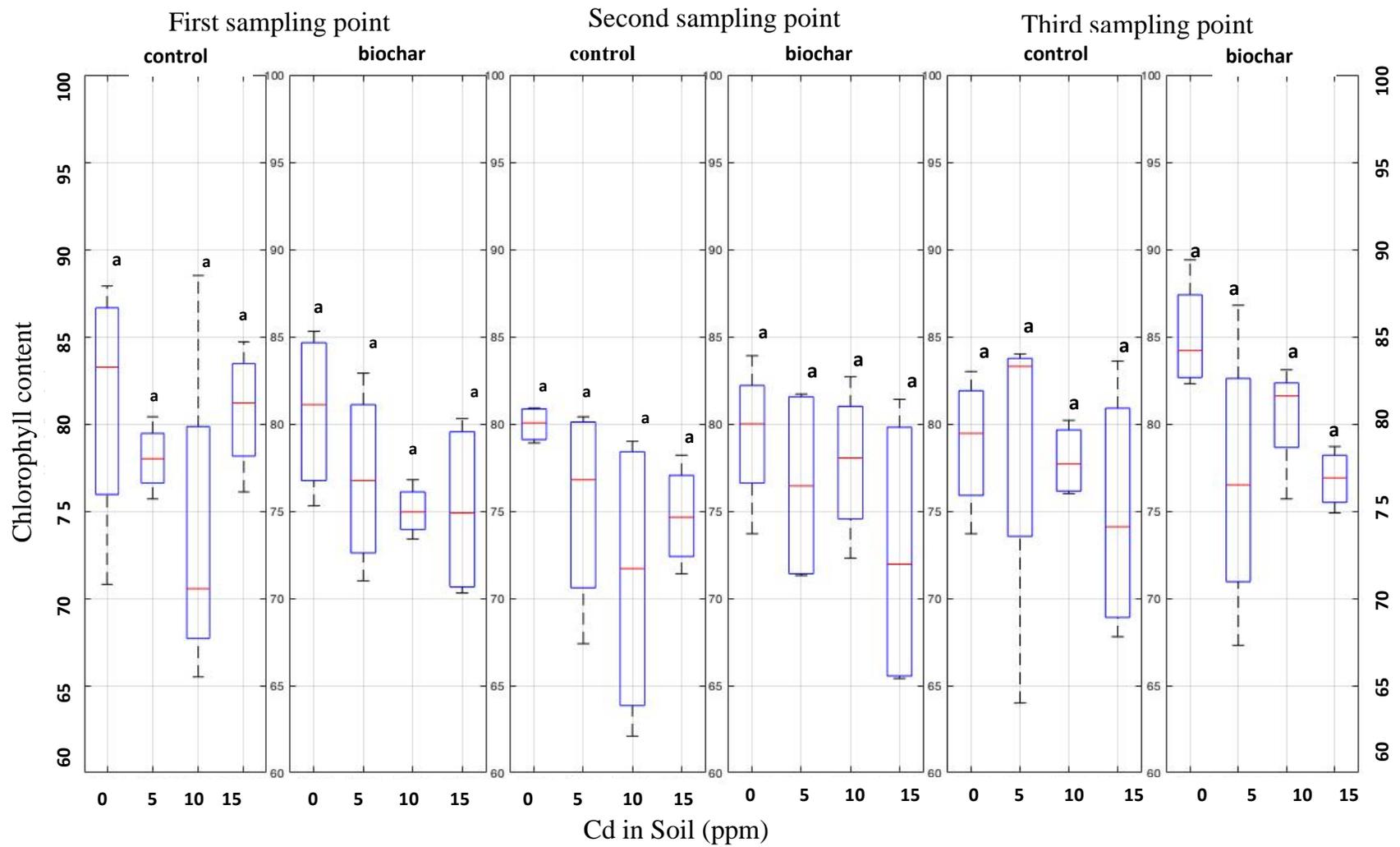


Figure 2.5 SPAD readings taken on the leaves of kale plants grown in soil amended with four Cd concentrations at three time points. Different letters represent significant differences between treatments ( $P < 0.05$ )

#### **2.4.2 Influence biochar addition on the growth, dry weight, elemental concentrations and SPAD readings of basil and kale plants subject to one of four soil Cd concentrations**

There were some visible differences in basil plants grown with and without biochar. Specifically, differences in flowering among basil plants subject to the different soil Cd levels were not as dramatic in plants amended with biochar compared to plants that did not receive biochar (Figure 2.1). Moreover, while there were no significant differences in plant height among basil plants treated or not with biochar and subject to individual Cd levels at any time point, there were subtle differences over time between plants with addition or not of biochar, with slower rate of growth in biochar amended pots (Figure 2.4). For example, during the second and third sampling point, in the absence of the biochar addition, basil plants subject to 5 ppm Cd were significantly taller than the control (0 ppm Cd rate), but there were no differences among soil Cd levels in plants that did not receive the biochar addition. There were no visible differences in development (Figure 2.1), or plant height among kale plants amended with or without, and subject to the different soil Cd concentrations at any time point (figure 2.7).

There were no differences in the aboveground dry plant biomass, total C and N, C:N ratio or Zn concentrations at harvest in basil or kale plants amended or not with biochar, and subject to the four soil Cd concentrations (Table 2.3). There were not significant differences in either basil and kale plants with respect to total Cd concentrations in the aboveground biomass, however, there were subtle differences between the two species (Table 2.3). In basil, increasing Cd concentrations in the aboveground biomass were correlated with increasing of soil Cd, while in kale, there was no difference between plants subject to 5 and 10 ppm Cd.

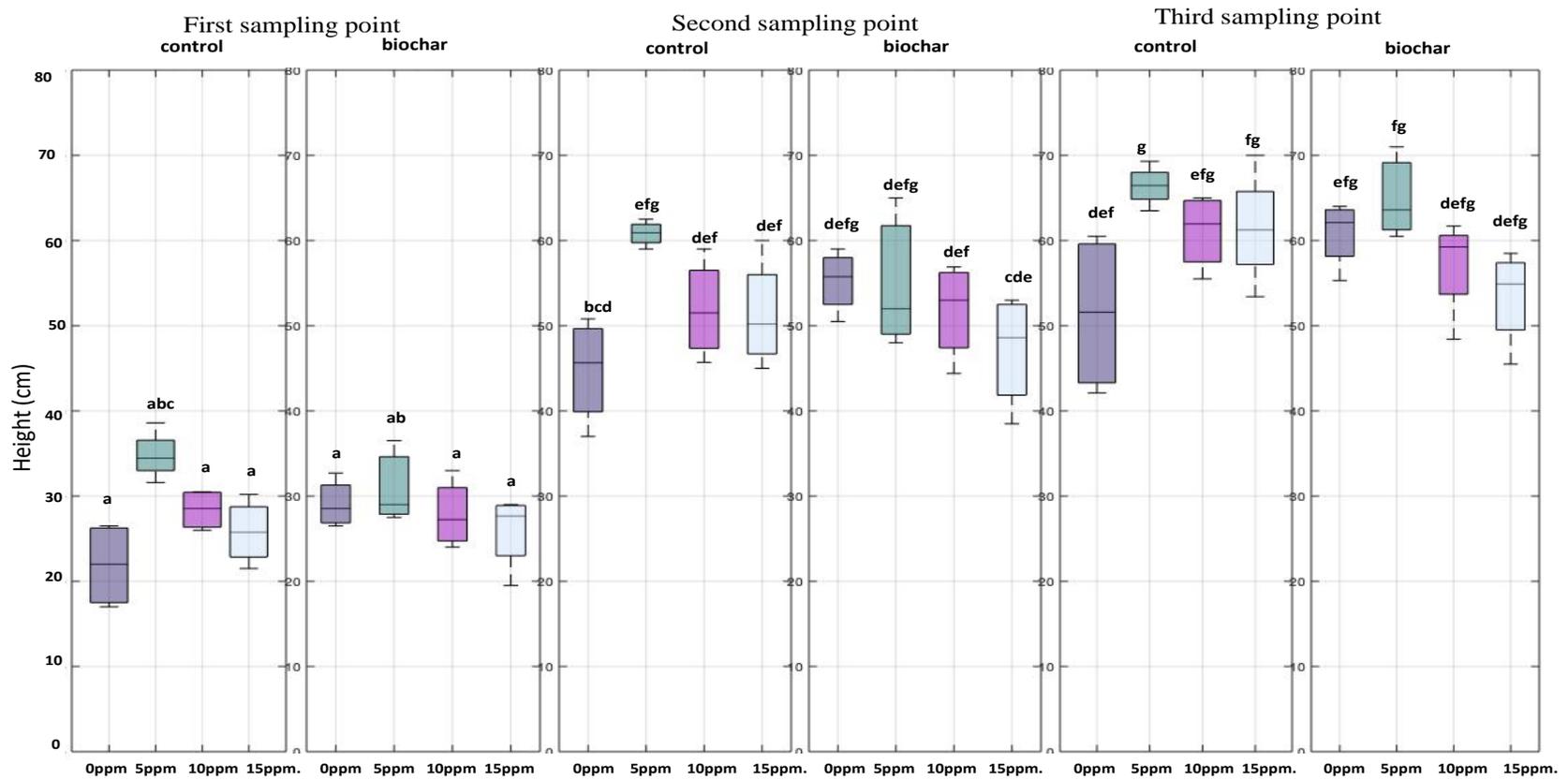


Figure 2.6 Height of basil plants grown with and without biochar (3% v/v) and subject to one of four soil Cd concentrations during three time points. Black bars represent basil plants without biochar (NB) and blue bars represent plants with biochar. Different letters represent significant differences between treatment ( $p < 0.05$ )

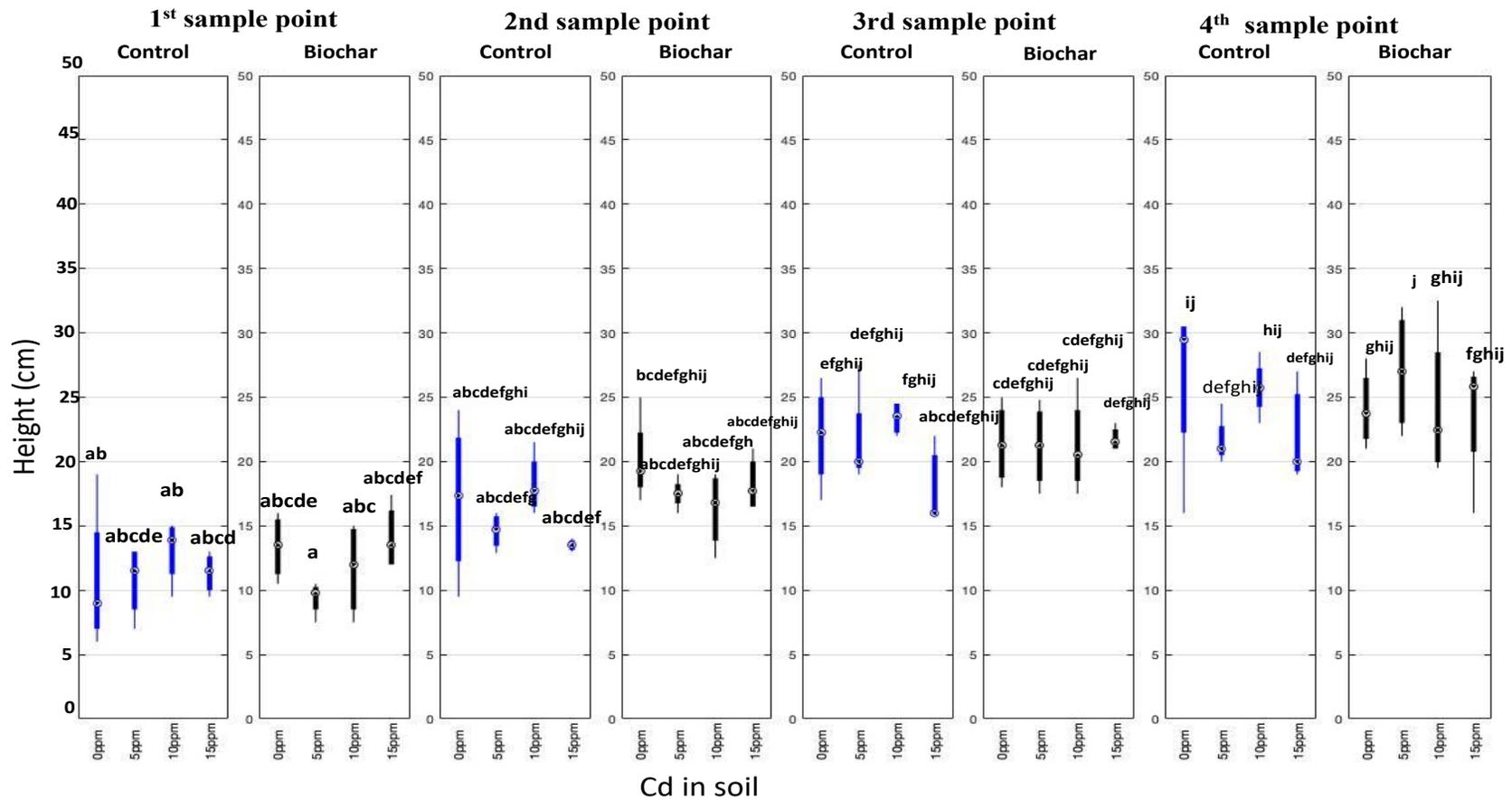


Figure 2.7 Height of kale plants grown with and without biochar (3% v/v) at one of four soil Cd concentrations during four time points. Black bars represent basil plants without biochar (NB) and blue bars represent plants with biochar. Different letter represents significant differences between treatment ( $p < 0.05$ )

### **2.4.3 Effect of soil Cd concentrations on plant physiological indices generating using HSI taken from the side and top view in basil plants**

Significant differences among plant vegetation indices were only apparent during the first sampling point regardless of whether images were acquired from the side or top view of basil plants (Tables 2.4 and 2.5). When vegetation indices were generated using the side view spectra during the first sampling point, NDVI, PSRI, NCPI and ARI were significantly different (Table 2.4). In addition, the index NCPI index detected differences among plants with/without the biochar addition, and NDVI, PSRI and ARI indices detected differences among basil plants with respect to the different Cd levels (Table 2.4). The PSRI index also detected an interaction between the biochar addition and Cd levels, indicating that it could detect subtle differences in Cd levels when plants had or not biochar addition

More indices were able to detect differences in basil when plants were imaged from the top view (Table 2.5). In particular, during the first sampling point, NDVI, NBNDVI, NRI, SIPI, PSRI, NDVI\_RE, CI\_RE, MSR\_RE and HMSSI were all significantly different (Table 2.5). Like the side view, most of these indices detected differences among the Cd levels, and only the PSRI index detected differences among plants amended with an addition or not of biochar (Table 2.

Table 2.4 Effect of soil amendment and soil cadmium concentration on fourteen hyperspectral indices generated using the side view of images of basil plants collected at three time points.

Soil amendment	Cd level	NDVI	NBNDVI	PRI	NRI	TCARI	SIPI	PSRI	PhRI	NPCI	ARI	NDVI_RE	CI_RE	MSR_RE	HMSSI
<b>Time point 1</b>															
<b>Biochar</b>	<b>0</b>	0.76	0.81	0	0.43	0.08	0.75	-	0.07	-0.36	-1.85	0.5	2.08	1.01	-69.25
<b>Control</b>	<b>0</b>	0.77	0.82	0	0.43	0.08	0.75	-0.028	0.07	-0.36	-1.91	0.51	2.2	1.06	-79.05
<b>Biochar</b>	<b>5</b>	0.76	0.816	0	0.438	0.087	0.75	-0.027	0.07	-0.35	-1.77	0.501	2.145	1.032	-78.809
<b>Control</b>	<b>5</b>	0.77	0.83	0	0.44	0.09	0.77	-	0.07	-0.36	-1.71	0.51	2.25	1.07	-91.5
<b>Biochar</b>	<b>10</b>	0.77	0.82	0	0.43	0.09	0.76	0.0262	0.07	-0.35	-1.79	0.51	2.17	1.04	-84
<b>Control</b>	<b>10</b>	0.78	0.83	0	0.44	0.09	0.77	-0.024	0.07	-0.36	-1.82	0.52	2.31	1.1	-94.4
<b>Biochar</b>	<b>15</b>	0.75	0.81	0.01	0.44	0.09	0.75	-0.027	0.07	-0.34	-1.66	0.48	1.97	0.97	-71.13
<b>Control</b>	<b>15</b>	0.75	0.82	0	0.46	0.09	0.75	-0.032	0.07	-0.35	-1.58	0.47	1.87	0.94	-57.94
<b>F-test</b>		S	NS	NS	NS	NS	NS	S	NS	S	S	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS
<b>Cd level</b>		S	Ns	NS	NS	NS	NS	S	NS	NS	S	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS	NS	NS
<b>Time point 2</b>															
<b>Biochar</b>	<b>0</b>	0.76	0.82	0	0.46	0.08	0.77	-0.03	0.07	-0.36	-2.02	0.51	2.31	1.08	-91.67
<b>Control</b>	<b>0</b>	0.77	0.83	0	0.45	0.07	0.77	-0.02	0.07	-0.37	-2.27	0.53	2.52	1.15	-103
<b>Biochar</b>	<b>5</b>	0.71	0.8	0.01	0.46	0.09	0.75	-0.03	0.06	-0.32	-1.68	0.46	1.9	0.93	-68.08
<b>Control</b>	<b>5</b>	0.75	0.82	0	0.47	0.08	0.77	-0.03	0.07	-0.35	-1.84	0.5	2.17	1.03	-82.11
<b>Biochar</b>	<b>10</b>	0.77	0.83	0	0.46	0.08	0.78	-0.03	0.07	-0.37	-2.2	0.53	2.45	1.13	-97.17
<b>Control</b>	<b>10</b>	0.78	0.84	0	0.47	0.08	0.78	-0.02	0.07	-0.37	-2.11	0.53	2.46	1.14	-102.73
<b>Biochar</b>	<b>15</b>	0.76	0.82	0	0.46	0.09	0.77	-0.03	0.07	-0.36	-1.83	0.5	2.2	1.05	-86.19
<b>Control</b>	<b>15</b>	0.76	0.83	0	0.48	0.09	0.77	-0.03	0.07	-0.36	-1.8	0.5	2.1	1.02	-79.27
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd level</b>		NS	NS	S	NS	NS	NS	NS	NS	NS	NS	S	NS	S	S

Table 2.4 continued

Interaction		NS	NS	NS	NS	NS	NS	NS	NS						
<b>Time point 3</b>															
<b>Biochar</b>	<b>0</b>	0.73	0.8	0	0.44	0.08	0.75	-0.03	0.06	-0.33	-1.86	0.49	2.17	1.03	-87.87
<b>Control</b>	<b>0</b>	0.75	0.81	0	0.44	0.07	0.76	-0.03	0.06	-0.35	-2.06	0.51	2.38	1.1	-92.53
<b>Biochar</b>	<b>5</b>	0.73	0.80	0.01	0.45	0.08	0.75	-0.03	0.06	-0.33	-1.81	0.48	2.07	0.99	-75.93
<b>Control</b>	<b>5</b>	0.72	0.8	0.01	0.45	0.08	0.75	-0.03	0.06	-0.32	-1.7	0.47	2.02	0.97	-76.8
<b>Biochar</b>	<b>10</b>	0.74	0.81	0	0.44	0.07	0.76	-0.03	0.06	-0.34	-2.07	0.51	2.32	1.08	-88.62
<b>Control</b>	<b>10</b>	0.75	0.82	0	0.47	0.08	0.77	-0.02	0.07	-0.35	-2.06	0.51	2.36	1.09	-96.2
<b>Biochar</b>	<b>15</b>	0.74	0.81	0	0.45	0.08	0.76	-0.03	0.06	-0.34	-1.88	0.5	2.19	1.04	-84.97
<b>Control</b>	<b>15</b>	0.74	0.82	0	0.47	0.08	0.76	-0.03	0.06	-0.34	-1.75	0.49	2.04	0.99	-75.37
<b>F-test</b>		NS	NS	NS	Ns	NS	NS	NS	NS						
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS						
<b>Cd level</b>		NS	NS	NS	NS	NS	NS	NS	NS						
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	N						

\*S refers to significant differences ( $p < 0.05$ )

Table 2.5 Effect of biochar and soil cadmium concentration on fourteen hyperspectral indices generated using the top view of images of basil plants collected at three time points.

Soil amendment	Cd level	NDVI	NBNDVI	PRI	NRI	TCARI	SIPI	PSRI	PhRI	NPCI	ARI	NDVI_RE	CI_RE	MSR_RE	HMSSI
<b>Time point 1</b>															
<b>Biochar</b>	0	0.76	0.82	-0.03	0.44	0.20	0.76	-0.0292	0.07	-0.34	-0.47	0.46	1.79	0.91	-61.60
<b>Control</b>	0	0.77	0.82	-0.03	0.42	0.18	0.76	-0.0265	0.07	-0.34	-0.51	0.48	1.95	0.97	-74.14
<b>Biochar</b>	5	0.77	0.83	-0.02	0.45	0.21	0.77	-0.0274	0.07	-0.35	-0.39	0.47	1.87	0.94	-68.67
<b>Control</b>	5	0.77	0.83	-0.02	0.44	0.21	0.77	-0.0254	0.07	-0.34	-0.34	0.48	1.96	0.97	-77.38
<b>Biochar</b>	10	0.76	0.81	-0.02	0.43	0.21	0.76	-0.0280	0.07	-0.33	-0.38	0.46	1.81	0.91	-64.93
<b>Control</b>	10	0.77	0.82	-0.02	0.44	0.20	0.77	-0.0253	0.07	-0.34	-0.38	0.48	1.94	0.97	-76.98
<b>Biochar</b>	15	0.75	0.81	-0.02	0.45	0.21	0.75	-0.0302	0.07	-0.33	-0.40	0.45	1.68	0.86	-55.93
<b>Control</b>	15	0.75	0.82	-0.02	0.46	0.22	0.76	-0.0289	0.07	-0.33	-0.24	0.43	1.59	0.83	-55.81
<b>F-test</b>		S	S	NS	S	NS	S	S	NS	NS	NS	S	S	S	S
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS	NS	S
<b>Cd level</b>		S	S	NS	S	NS	S	S	NS	NS	NS	S	S	S	S
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Time point 2</b>															
<b>Biochar</b>	0	0.77	0.84	-0.03	0.48	0.20	0.79	-0.03	0.07	-0.34	-0.56	0.48	2.00	0.98	-76.18
<b>Control</b>	0	0.79	0.85	-0.03	0.47	0.18	0.80	-0.02	0.07	-0.36	-0.61	0.51	2.19	1.05	-88.29
<b>Biochar</b>	5	0.77	0.84	-0.02	0.49	0.21	0.79	-0.03	0.07	-0.35	-0.53	0.48	1.98	0.97	-75.01
<b>Control</b>	5	0.77	0.84	-0.02	0.50	0.20	0.79	-0.03	0.07	-0.35	-0.53	0.48	1.97	0.97	-75.26
<b>Biochar</b>	10	0.78	0.84	-0.03	0.48	0.19	0.79	-0.03	0.07	-0.35	-0.61	0.50	2.11	1.02	-82.28
<b>Control</b>	10	0.78	0.85	-0.03	0.49	0.20	0.80	-0.03	0.07	-0.36	-0.55	0.50	2.10	1.02	-82.33
<b>Biochar</b>	15	0.77	0.84	-0.02	0.49	0.21	0.79	-0.03	0.07	-0.35	-0.49	0.48	1.94	0.96	-71.64
<b>Control</b>	15	0.77	0.84	-0.03	0.49	0.22	0.79	-0.03	0.07	-0.35	-0.43	0.47	1.87	0.94	-68.66
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd level</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Time point 3</b>															
<b>Biochar</b>	0	0.75	0.82	-0.02	0.47	0.18	0.77	-0.03	0.06	-0.33	-0.57	0.48	1.99	0.97	-76.08

Table 2.5 continued

<b>Control</b>	0	0.76	0.83	-0.02	0.46	0.17	0.78	-0.03	0.06	-0.35	-0.64	0.50	2.20	1.04	-86.66
<b>Biochar</b>	5	0.75	0.83	-0.02	0.48	0.19	0.78	-0.03	0.06	-0.34	-0.53	0.47	1.98	0.96	-72.40
<b>Control</b>	5	0.73	0.81	-0.02	0.47	0.19	0.77	-0.03	0.06	-0.32	-0.50	0.46	1.88	0.93	-69.29
<b>Biochar</b>	10	0.76	0.83	-0.02	0.46	0.18	0.78	-0.03	0.06	-0.33	-0.66	0.49	2.14	1.02	-79.58
<b>Control</b>	10	0.76	0.84	-0.02	0.48	0.19	0.79	-0.03	0.06	-0.34	-0.61	0.49	2.11	1.01	-80.35
<b>Biochar</b>	15	0.76	0.83	-0.02	0.48	0.19	0.78	-0.03	0.06	-0.34	-0.57	0.48	2.03	0.98	-76.24
<b>Control</b>	15	0.77	0.84	-0.02	0.49	0.20	0.79	-0.03	0.07	-0.34	-0.52	0.48	1.96	0.97	-75.37
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd level</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

\*S refers to significant differences ( $p < 0.05$ )

#### **2.4.4 Effect of soil Cd concentrations on plant physiological indices generating using HSI taken from the side and top view in kal plants**

Like basil plants, differences among plant physiological conditions in kale plants were only apparent during the first sampling point, but fewer indices were able to detect differences due to biochar addition or Cd levels (Tables 2.6 and 2.7). When images were taken from the side view, the overall model was significant for the TCARI index, but neither soil amendment or Cd levels were significant (Table 2.6). However, there was an interaction, indicating that there could be subtle differences among individual treatments. The ARI index was also significantly different with respect to the biochar addition and there was also an interaction, indicating there could be subtle differences among Cd levels (Table 2.6).

When images were taken from the top view, ARI and CI\_RE were significant (Table 2.7). The ARI index detected differences with respect to biochar addition, soil Cd rate and their interaction, indicating that this is a powerful index for detecting differences in kale to Cd stress. The CI\_RE index was able to detect differences in Cd rate (Table 2.7).

Table 2.6 Effect of soil amendment and soil cadmium concentration on fourteen hyperspectral indices generated using the side view of images of kale plants collected at three time point.

Soil amendment	Cd level	NDVI	NBNDVI	PRI	NRI	TCARI	SIPI	PSRI	PhRI	NPCI	ARI	NDVI_RE	CI_RE	MSR_RE	HMSSI
<b>Time point 1</b>															
<b>Biochar</b>	<b>0</b>	0.71	0.74	-0.02	0.2	0.02	0.65	-0.05	0.02	-0.34	-2.3	0.6	3.26	1.39	-64.58
<b>Control</b>	<b>0</b>	0.66	0.7	-0.02	0.21	0.03	0.6	-0.06	0.02	-0.31	-1.28	0.52	2.34	1.1	-42.69
<b>Biochar</b>	<b>5</b>	0.7	0.73	-0.02	0.2	0.02	0.63	-0.05	0.02	-0.34	-2.1	0.57	2.87	1.27	-54.63
<b>Control</b>	<b>5</b>	0.69	0.72	-0.02	0.2	0.03	0.63	-0.05	0.02	-0.33	-2.06	0.56	2.78	1.24	-53.46
<b>Biochar</b>	<b>10</b>	0.68	0.71	-0.02	0.19	0.02	0.61	-0.06	0.02	-0.33	-2.14	0.56	2.79	1.25	-47.66
<b>Control</b>	<b>10</b>	0.68	0.71	-0.02	0.21	0.03	0.62	-0.05	0.02	-0.32	-1.72	0.55	2.67	1.21	-51.24
<b>Biochar</b>	<b>15</b>	0.69	0.72	-0.02	0.21	0.03	0.63	-0.05	0.02	-0.33	-1.81	0.55	2.68	1.21	-52.36
<b>Control</b>	<b>15</b>	0.69	0.72	-0.02	0.19	0.02	0.62	-0.06	0.02	-0.33	-2.24	0.57	2.79	1.25	-49.23
<b>F-test</b>		NS	NS	NS	NS	S	NS	NS	NS	NS	S	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS
<b>Cd level</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	S	NS	S	NS	NS	NS	NS	S	NS	S	S	NS
<b>Time point 2</b>															
<b>Biochar</b>	<b>0</b>	0.71	0.74	-0.05	0.18	0.02	0.63	-0.07	0.01	-0.35	-1.5	0.61	3.45	1.44	-49.14
<b>Control</b>	<b>0</b>	0.7	0.73	-0.04	0.19	0.04	0.62	-0.07	0.01	-0.34	-1.13	0.58	3.01	1.31	-42.73
<b>Biochar</b>	<b>5</b>	0.71	0.74	-0.05	0.18	0.03	0.63	-0.07	0.01	-0.35	-1.29	0.6	3.18	1.37	-43.36
<b>Control</b>	<b>5</b>	0.69	0.72	-0.04	0.17	0.03	0.61	-0.07	0.01	-0.33	-1.22	0.58	2.97	1.31	-39.9
<b>Biochar</b>	<b>10</b>	0.69	0.71	-0.05	0.17	0.02	0.6	-0.08	0	-0.34	-1.35	0.59	3.06	1.33	-38.1
<b>Control</b>	<b>10</b>	0.7	0.73	-0.04	0.19	0.04	0.62	-0.07	0.01	-0.34	-1.17	0.58	3.08	1.33	-42.28
<b>Biochar</b>	<b>15</b>	0.69	0.73	-0.04	0.19	0.04	0.62	-0.07	0.01	-0.34	-1.18	0.57	2.95	1.29	-39.7
<b>Control</b>	<b>15</b>	0.68	0.71	-0.04	0.18	0.02	0.6	-0.08	0	-0.33	-1.25	0.58	2.91	1.29	-36.63
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amenment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd level</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Time point 3</b>															
<b>Biochar</b>	<b>0</b>	0.69	0.72	-0.04	0.18	0.04	0.61	-0.07	0.01	-0.33	-1.13	0.57	3.02	1.3	-41.31

Table 2.6 continued

<b>Control</b>	<b>0</b>	0.67	0.71	-0.04	0.19	0.05	0.61	-0.07	0.01	-0.31	-0.85	0.54	2.71	1.2	-39.2
<b>Biochar</b>	<b>5</b>	0.68	0.71	-0.04	0.17	0.03	0.6	-0.08	0	-0.32	-1.11	0.57	2.97	1.29	-38.74
<b>Control</b>	<b>5</b>	0.67	0.7	-0.04	0.16	0.02	0.59	-0.08	0	-0.31	-1.16	0.57	2.94	1.29	-38.35
<b>Biochar</b>	<b>10</b>	0.65	0.69	-0.04	0.16	0.02	0.57	-0.09	0	-0.31	-1.12	0.56	2.8	1.24	-33.21
<b>Control</b>	<b>10</b>	0.67	0.71	-0.04	0.18	0.04	0.6	-0.07	0.01	-0.32	-1	0.56	2.88	1.25	-38.57
<b>Biochar</b>	<b>15</b>	0.67	0.71	-0.04	0.18	0.03	0.59	-0.08	0.01	-0.32	-1.03	0.56	2.77	1.23	-34.93
<b>Control</b>	<b>15</b>	0.66	0.69	-0.04	0.17	0.01	0.58	-0.08	0	-0.31	-1.16	0.56	2.82	1.25	-34.14
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd level</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

\*S refers to significant differences ( $p < 0.05$ )

Table 2.7 Effect of soil amendment and soil cadmium concentration on fourteen hyperspectral indices generated using the top view of images of kale plants collected at three time points

Soil amendment	Cd level	NDVI	NBNDVI	PRI	NRI	TCARI	SIPI	PSRI	PhRI	NPCI	ARI	NDVI_RE	CI_RE	MSR_RE	HMSSI
<b>Time point 1</b>															
<b>Biochar</b>	<b>0</b>	0.75	0.77	-0.04	0.17	0.03	0.68	-0.05	0.01	-0.35	-1.37	0.63	3.72	1.53	-68.36
<b>Control</b>	<b>0</b>	0.73	0.76	-0.04	0.19	0.05	0.66	-0.05	0.02	-0.33	-0.77	0.59	3.01	1.33	-57.48
<b>Biochar</b>	<b>5</b>	0.74	0.76	-0.04	0.19	0.06	0.67	-0.05	0.02	-0.33	-0.84	0.59	3.06	1.34	-58.89
<b>Control</b>	<b>5</b>	0.73	0.76	-0.04	0.20	0.07	0.67	-0.05	0.02	-0.33	-0.79	0.58	2.96	1.30	-57.31
<b>Biochar</b>	<b>10</b>	0.73	0.75	-0.04	0.18	0.05	0.66	-0.06	0.01	-0.33	-0.86	0.59	3.08	1.35	-56.29
<b>Control</b>	<b>10</b>	0.74	0.76	-0.04	0.19	0.06	0.68	-0.05	0.02	-0.33	-0.65	0.59	3.10	1.35	-63.47
<b>Biochar</b>	<b>15</b>	0.74	0.77	-0.04	0.20	0.06	0.68	-0.05	0.02	-0.34	-0.82	0.60	3.10	1.35	-63.70
<b>Control</b>	<b>15</b>	0.72	0.75	-0.04	0.18	0.05	0.65	-0.06	0.01	-0.33	-0.86	0.59	3.02	1.33	-53.73
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	S	NS	S	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS
<b>Cd rate</b>		NS	NS	S	NS	NS	NS	NS	NS	NS	S	NS	S	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS
<b>Time point 2</b>															
<b>Biochar</b>	<b>0</b>	0.71	0.74	-0.05	0.18	0.02	0.63	-0.07	0.01	-0.35	-1.50	0.61	3.45	1.44	-49.14
<b>Control</b>	<b>0</b>	0.70	0.73	-0.04	0.19	0.04	0.62	-0.07	0.01	-0.34	-1.13	0.58	3.01	1.31	-42.73
<b>Biochar</b>	<b>5</b>	0.71	0.74	-0.05	0.18	0.03	0.63	-0.07	0.01	-0.35	-1.29	0.60	3.18	1.37	-43.36
<b>Control</b>	<b>5</b>	0.69	0.72	-0.04	0.17	0.03	0.61	-0.07	0.01	-0.33	-1.22	0.58	2.97	1.31	-39.90
<b>Biochar</b>	<b>10</b>	0.69	0.71	-0.05	0.17	0.02	0.60	-0.08	0.00	-0.34	-1.35	0.59	3.06	1.33	-38.10
<b>Control</b>	<b>10</b>	0.70	0.73	-0.04	0.19	0.04	0.62	-0.07	0.01	-0.34	-1.17	0.58	3.08	1.33	-42.28
<b>Biochar</b>	<b>15</b>	0.69	0.73	-0.04	0.19	0.04	0.62	-0.07	0.01	-0.34	-1.18	0.57	2.95	1.29	-39.70
<b>Control</b>	<b>15</b>	0.68	0.71	-0.04	0.18	0.02	0.60	-0.08	0.00	-0.33	-1.25	0.58	2.91	1.29	-36.63
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd rate</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

Table 2.7 continued

Time point 3															
<b>Biochar</b>	<b>0</b>	0.69	0.72	-0.04	0.18	0.04	0.61	-0.07	0.01	-0.33	-1.13	0.57	3.02	1.30	-41.31
<b>Control</b>	<b>0</b>	0.67	0.71	-0.04	0.19	0.05	0.61	-0.07	0.01	-0.31	-0.85	0.54	2.71	1.20	-39.20
<b>Biochar</b>	<b>5</b>	0.68	0.71	-0.04	0.17	0.03	0.60	-0.08	0.00	-0.32	-1.11	0.57	2.97	1.29	-38.74
<b>Control</b>	<b>5</b>	0.67	0.70	-0.04	0.16	0.02	0.59	-0.08	0.00	-0.31	-1.16	0.57	2.94	1.29	-38.35
<b>Biochar</b>	<b>10</b>	0.65	0.69	-0.04	0.16	0.02	0.57	-0.09	0.00	-0.31	-1.12	0.56	2.80	1.24	-33.21
<b>Control</b>	<b>10</b>	0.67	0.71	-0.04	0.18	0.04	0.60	-0.07	0.01	-0.32	-1.00	0.56	2.88	1.25	-38.57
<b>Biochar</b>	<b>15</b>	0.67	0.71	-0.04	0.18	0.03	0.59	-0.08	0.01	-0.32	-1.03	0.56	2.77	1.23	-34.93
<b>Control</b>	<b>15</b>	0.66	0.69	-0.04	0.17	0.01	0.58	-0.08	0.00	-0.31	-1.16	0.56	2.82	1.25	-34.14
<b>F-test</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Soil amendment</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Cd rate</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
<b>Interaction</b>		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

\*S refers to significant differences (p<0.05)

#### **2.4.5 Effect of soil Cd concentration on colormaps and reflectance graphs generated using NDVI data collected during the first and third sampling time points**

The individual color maps produced from basil and kale plants were performed to visualize the vegetation index distribution on the plant. Each pixel of these images corresponds to the NDVI value represented in a color. With an NDVI close to 1 the pixel is redder and is correlated with more chlorophyll content, while the closer the pixel is to 0 it appears as bluer. This allows visualization of areas where the NDVI differs in the plant. Normal range for NDVI is between -1 to 1, but in this study values fell within 0 to 1. In plants grown in soil with biochar addition and subject to 0 and 15 ppm soil Cd concentrations during the first sampling point did not indicate that there were clear differences in NDVI between treatments (Figure 2.8). However, during the third sampling point, differences between plants grown in the two soil Cd treatments were detectable between old and new leaves. In particular, newer basil leaves in the 0 ppm soil Cd treatment had a higher NDVI value than that of the older leaves found on the periphery, whereas newer leaves in 15 ppm Cd treatment had a lower NDVI value than that of the older leaves found in the periphery. The greater number of flowers in the 0 relative to 15 soil Cd treatment, were also apparent during the third sampling point. When comparing the reflectance spectra, there were no differences between the first and final sampling point in basil plants subject to 0 ppm Cd, though there were differences between the two time points in plants subjected to soil Cd concentrations of 15 ppm (Figure 2.8). Specifically, during the first sampling time point in the 15 ppm treatment, the NDVI was lower, which is related to a decrease in chlorophyll content but also to more light reflected, while the NDVI reflectance was higher at the final sampling point.

Like basil, there did not appear to be any clear differences in the individual color maps produced from kale plants grown in soil amended with biochar and subject to 0 and 15 ppm soil Cd concentrations during the first or final sampling point (Figure 2.9). However, during the third sampling point, each Cd treatment plant differed in the reflectance between the leaf tips and the center of the plant. The tips had a bluer and greenish color correlated with less chlorophyll. There were also differences in the reflectance spectra (Figure 2.9). Specifically, in the plant subject to 0 ppm soil Cd, the reflectance was lower during the first relative to the third week, while in the plant grown in 15 ppm Cd, there were differences in lower wavelengths between the first and third sampling point (Figure 2.9).

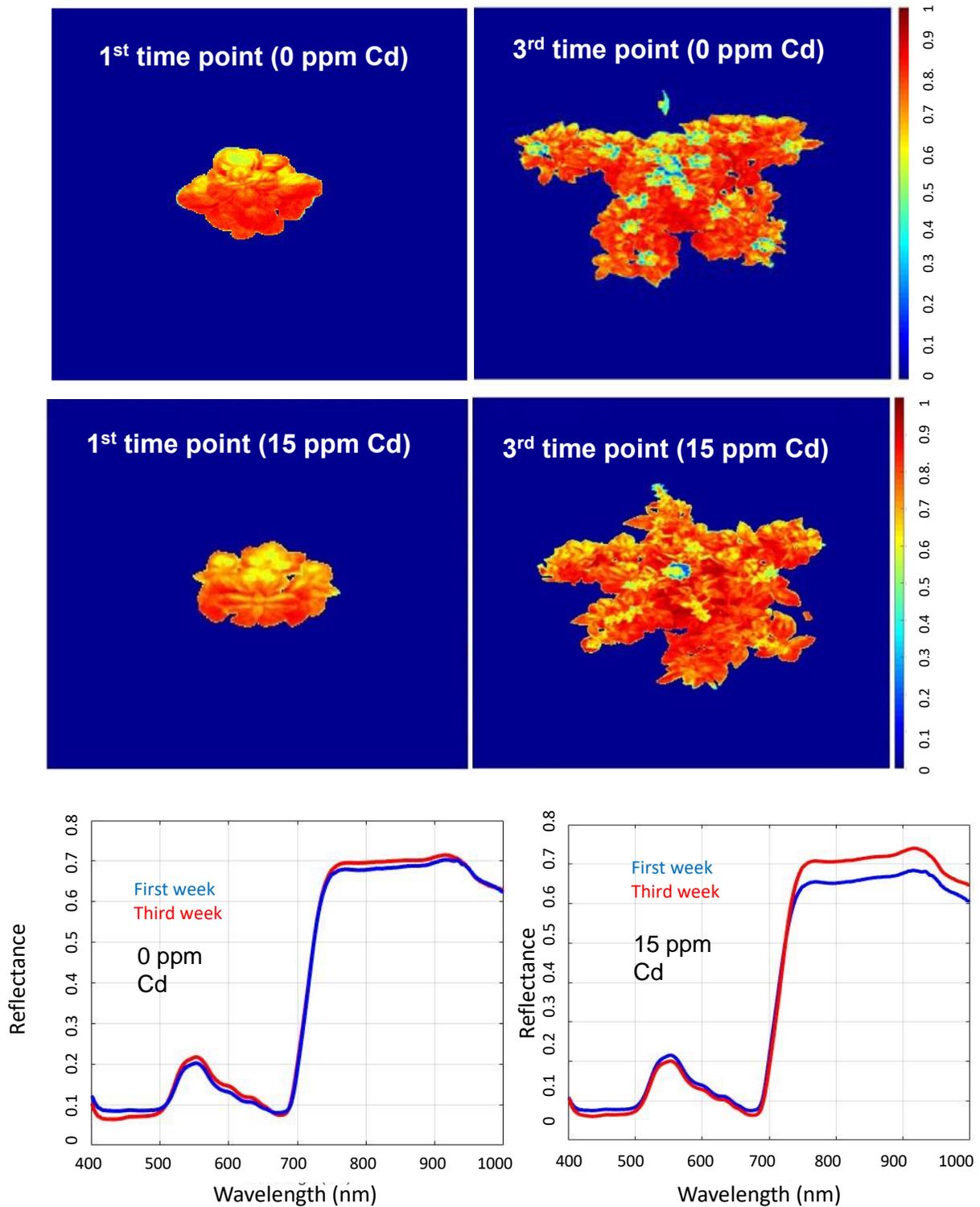


Figure 2.8 NDVI color maps and reflectance graphs of basil plants grown in soil amended with biochar and subject to 0 or 15 ppm soil Cd during the first and third sampling points.

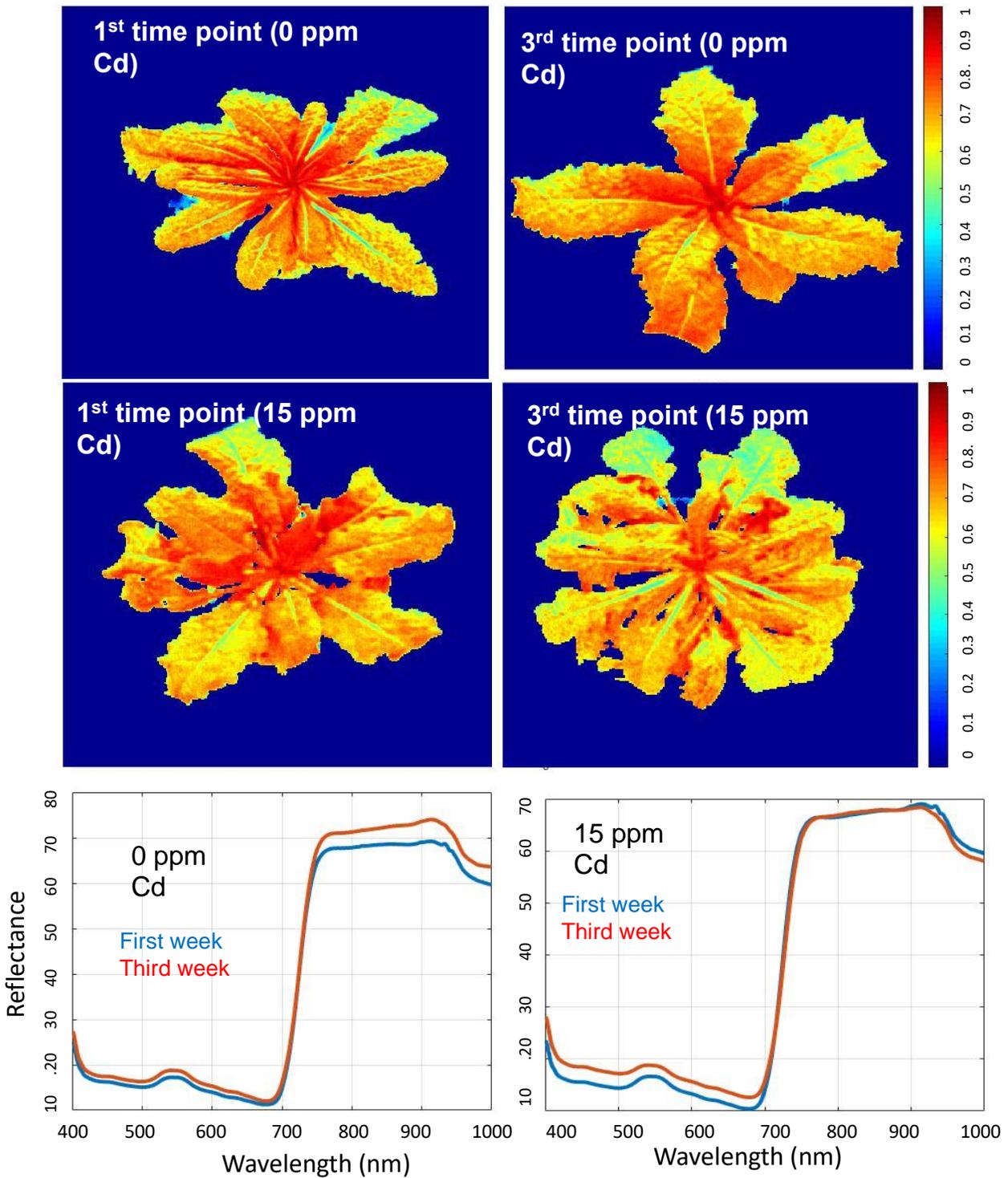


Figure 2.9 NDVI color maps and reflectance graphs of kale plants grown in soil amended with biochar and subject to 0 or 15 ppm soil Cd during the first and third sampling point.

## 2.5 Discussion

Practical approaches to immobilize Cd in soil and prevent uptake into edible plant tissues are needed to improve crop production and protect human health. Current post-harvest approaches used to quantify Cd uptake in plants is time consuming and expensive, which makes identification of effective remediation strategies difficult. Consequently, the primary purpose of this study was to determine if HSI can be used to rapidly detect Cd stress and predict Cd uptake during crop production. To answer this question, we conducted experiments using two distinct leafy green species that are known to accumulate Cd, yet generally differ in their physiological responses to Cd stress. We also aimed to determine whether a locally sourced biochar addition could increase plant health and prevent Cd stress in these crops, and if subtle differences in plant responses to the presence of this soil amendment could be detected by HSI.

As predicted, visible differences in plant morphology due to soil Cd stress were visually observable in basil, but not kale plants (Figs. 2.1, 2.2, 2.3). In particular, basil plants subject to 5 ppm soil Cd appeared to flower earlier and were significantly taller than the control (0 ppm soil Cd), and growth rates in basil over time in all Cd treatments were greater than the control (Fig. 2.1; Fig. 2.2). Similar developmental responses to low levels of Cd stress have been observed in other leafy green plants (Baldantoni et al., 2016). We suspect that this could be due to a phenomenon known as hormesis, where an organism exhibits a biphasic response to low amounts of a toxic substance such as heavy metals (Morkunas et al., 2018). This overcompensation in plant growth is theorized to be an evolutionary adaptation to low levels of plant stress (Calabrese et al., 2015). Based on these results, it is possible that basil plants could be used as a quick, simple test to determine if remediation strategies are working in the field, though additional studies are needed to confirm this hypothesis.

As expected, we observed no visible stress responses in kale plants (Fig. 2.1; Fig. 2.3) despite the fact that kale plants accumulated more total Cd in their aboveground biomass than basil (Table 2.2). The lack of a visual response to increasing concentrations of soil Cd corresponds with the results of Jakovljevc et al. (2013), who also observed no visible responses in several Brassica species including kale, to soil Cd levels of 5 and 10 ppm. In another experiment comparing levels of 4, 8 and 16 ppm soil Cd showed differences in plant biomass in kale at the highest soil concentration (Haghighi et al. 2016). Therefore, it is possible that Cd levels in our study were not high enough to cause visually observable plant response. The results of this study confirm what

previous studies have indicated that Brassica species like kale could be used in phytoremediation efforts to extract Cd from soil (Mourato et al., 2015). However, they also provide additional evidence to indicate that growing these crops on soils that could be contaminated by heavy metals poses high risks for human health.

Biochar has received a lot of interest in recent years for its potential to improve crop productivity (Shoaf et al., 2016), as well as immobilize soil Cd and prevent the uptake of this heavy metal into edible plant tissues (Kim et al., 2015; Gomez and Hoagland, unpublished). In this study, visible differences in flowering and plant growth among basil plants subject to soil Cd were not apparent in plants with biochar addition, in comparison to plant that did not receive this amendment (Figures 2.1 and 21.). This suggests that the biochar evaluated in this study did reduce Cd stress in this sensitive plant species, albeit the effects were subtle. However, the biochar amendment did not significantly reduce Cd uptake in either basil or kale (Table 2.2). The inability of this amendment to prevent Cd uptake could be due to the fact that this particular biochar is not effective at immobilizing soil Cd, or that higher rate of this particular amendment would be needed to achieve the desired result. The lack of a response could also be due to the type of soil used in this study.

Given subtle differences in plant responses to the presence of toxic heavy metals like Cd, non-destructive tools are needed to quantify Cd stress and estimate uptake. As described above, Cd can interfere with chlorophyll synthesis and N dynamics in crops (He et al., 2015). Consequently, SPAD meters, which are commonly used to estimate chlorophyll concentrations in plants, are one tool that could potentially be used to help detect Cd stress. However, this was not the case in this study. There were some subtle differences in SPAD readings in basil plants subject to different soil Cd levels during different sampling points (figure 2.4), but we did not observe any significant differences in SPAD readings in kale (figure 2.5). Thus, we conclude that SPAD meters could provide some limited value for detecting Cd stress in basil, but not kale.

Vegetation indices (VIs) derived from a range of wavelengths have also been used to detect plant stress and estimate concentrations of various compounds in plants such as chlorophyll, N, anthocyanins, water and heavy metals (Ashourloo et al., 2016). Many of these indices have integrated red-edge information and narrow-band spectral data ( $\leq 10$  nm) to more accurately detect changes in plant physiology caused by many plant stresses (Zhang et al. 2018), but relatively little is about how these indices differ in their potential to detect heavy metal stress. Consequently, we

investigated whether fourteen plant physiological indices could be used to detect Cd stress, and quantify subtle differences in plant physiology due to a biochar soil amendment in our experiment with basil and kale plants. In addition, we aimed to determine if taking images from the top or side view of plants would affect the potential of these images to detect Cd stress and the effects of the biochar amendment.

One of the most commonly used VIs is the normalized difference vegetation index (NDVI). NDVI is sensitive to photosynthetic activity and is used to measure the greenness of plants (Sridhar et al., 2007; Tatishvili et al., 2018.). This index has previously been shown to be sensitive to chlorophyll concentration and cellular structure in plants (Sridhar et al., 2007), and some have suggested that it could be helpful in identifying symptoms of heavy metal toxicity (Yi 2019; Zhou et al., 2018). This is because Cd has been shown to inhibit photosynthesis and alter plant chlorophyll concentrations, even in plants that are considered hyperaccumulators (Zhou and Qiu, 2005). Chlorophyll absorbs visible light from 0.4 to 0.7  $\mu\text{m}$  (Weier & Herring , 2000) Cell structures in leaves reflect near-infrared light from 0.7 to 1.1  $\mu\text{m}$  (Weier & Herring , 2000). Consequently, stressed vegetation tends to reflect more red light and less near-infrared light, because photosynthetic processes absorb most of the red light and the near infrared light is reflected (Tatishvili et al., 2018) The more leaves a plant has, the more these wavelengths tend to be reflected. Calculations of NDVI result in a number that ranges from -1 to +1. A value close to +1 indicates the highest possible density of green leaves, though values around 0.7 represent the average value in most non-stressed plants (Weier & Herring , 2000)

Within the basil plants evaluated in this study, NDVI was able to detect differences caused by soil Cd, regardless of whether images were taken from the side or top or side view of the plants (Tables 2. 4 and 2. 5), validating previous reports that NDVI can be used to detect Cd stress in some plant species (Zhou et al., 2018). The pixel-level “heatmap” and reflectance spectra shown in Figure 2.8, illustrate how NDVI can be used to detect Cd stress in basil. For example, while all plants maintained reflectance near 0.70, subtle differences were visible between plants subject to 0 and 15 ppm soil Cd. However, while some subtle differences in NDVI profiles were apparent when comparing the individual colormaps and reflectance spectra in kale (Figure 2.9), this index was not able to detect overall treatment differences (Tables 2.6 and 2. 7), indicating that it will not work in all crop plants. The lack of response in kale relative to basil could be related to the lower overall N concentrations in kale relative to basil plants (Table 2.2).

Some have argued that NDVI has challenges related to saturation, which prevents this index from being able to accurately estimate biomass, and that the narrow band normalized difference vegetation index (NBNDVI) can be used to overcome this problem (Mutanga and Skidmore, 2004). In this study, NBNDVI was able to detect Cd stress in basil when images were taken from the top view (Table 2.5), but not from the side view (Table 2.4), or with either view in kale (Tables 2.6 and 2.7). Consequently, we conclude that NBNDVI is not a reliable indicator for quantifying Cd stress in leafy green crops

Other researchers have suggested that red-edge approaches could help improve the accuracy of plant physiological indices for detecting heavy metal stress in crops (Zhang et al., 2018). For example, the normalized difference vegetation at red edge (NDVI\_RE), chlorophyll index at red edge (CI\_RE) and modified simple ratio at red edge (MSR\_RE) have been developed to linearize the relationship between indices and biophysical parameters at the red edge. However, while models that include red edge indices are often sensitive to chlorosis and reductions in plant biomass, they are not susceptible to changes in internal cellular structure and therefore must be used with other indices to detect stress in plants (Sridhar et al., 2007). In this study, all three of these red edge indices were able to detect Cd stress in basil when the images were taken from the top view (Table 2.5), and the CI\_RE index was also able to detect Cd stress in kale when plants were imaged from the top view (Tables 2.7). Consequently, results of this study confirm that by coupling standard indices with those that include red edges, researchers may be able to detect more subtle changes in the responses of leafy green plants to Cd stress.

Because of the strong relationship between chlorophyll concentration and plant stress due to the presence of heavy metals, other indices that quantify chlorophyll concentrations such as the nitrogen reflectance index (NRI), the normalized pigment chlorophyll index (NPCI), and the transformed chlorophyll absorption reflectance index (TCARI)(Huang et al., 2014) might be useful in leafy greens . NRI and NPCI have previously been used to quantify changes in plant N status (Huang et al., 2014). The TCARI index exhibited high sensitivity for quantifying the leaf area index (LAI) among plants grown in different soils in comparison with single red-edge indices ((Bandaru et al., 2016)) to suggest that it could be a strong index for quantifying heavy metal stress in plants (Zhang et. al 2018). In this study, NRI was able to detect differences in Cd stress among basil plants when images were taken from the top view (Table 2.5), but not when viewed from the side (Table 2.4), or when using either view in kale (Table 2.6 and 2.7), indicating that it is not

particularly sensitive in leafy green crops subject to Cd stress. The model for TCARI was significant for kale when plants were imaged from the side view (Table 2.6), but neither Cd stress or soil amendment were significant, indicating this is not a very sensitive measurement. Interestingly, NPCI was able to detect differences among basil plants when plants were imaged from the side view, but the index detected differences in the biochar amendment addition rather than Cd stress (Table 2.4). Consequently, the NPCI index might be able to detect other subtle differences in plant physiological status caused by biochar such as water status, or concentrations of other nutrients besides N.

Other indices that could be useful in detecting subtle changes in plant physiological status due to Cd stress or soil amendments additions like biochar, include those designed to detect differences in other plant pigments such as carotenoids. This is because carotenoids act as accessory light-harvesting pigments, and are also expected to play an essential photoprotective role in plants (Young, 1991). The structure insensitive pigment index (SIPI), is said to be sensitive to the ratio of chlorophyll to bulk carotenoids in plants while minimizing the impact of the variable canopy structure, thereby potentially making it a valuable indicator for quantify plant stress ((Yu et al., 2018)). Similarly, the plant senescence reflectance index (PSRI) was also designed to quantify carotenoid pigments, and therefore could also be useful in detecting plant stress caused by heavy metals (Zhang et al., 2018). In this study, the SIPI index was able to detect Cd stress in basil when images were taken from the top view (Table 2.5), but not when they were taken from the side view (Table 2.4), or from either view in kale (Tables 2.6 and 2.7). Consequently, we conclude that the SIPI index is not particularly valuable in detecting Cd stress in leafy green crops. In contrast, PSRI was able to detect Cd stress in basil when images were taken from the side view and there was an interaction with the soil biochar addition (Table 2.4), indicating that this index could also detect subtle changes in plant physiological status caused by the biochar addition. Moreover, the PSRI index also detected changes caused by soil Cd and biochar when basil plants were imaged from the top view (Table 2.5), indicating that this index is powerful for detecting subtle changes in plant physiological status in basil.

Indices that quantify ratios between chlorophyll and anthocyanins, such as the anthocyanin reflectance index (ARI), could also be useful in the context of this study, as concentrations of anthocyanins have also been demonstrated to play a protective role in plant stress due to heavy metals (Baek et al., 2012). Results of this study verify the strong potential of the ARI index, as it

was able to detect Cd stress in basil when plants were imaged from the side view (Table 2.4), and both Cd stress and biochar amendments in kale when plants were imaged from both the top and side view (Tables 2.6 and 2.7). Consequently, we conclude that ARI is a sensitive index for detecting subtle differences in plant physiological status in leafy green crops.

Finally, the heavy metal stress index (HMSSI) was specifically developed to detect heavy metal stress in plants (Zhang et al., 2018). This new vegetative index is based on two red-edge indices, PSRI and CI\_RE (Zarcotejada et al., 2005). When plant stress increases, the value of the CI\_RE is expected to decrease while the value of PSRI increases, making this index particularly valuable for detecting heavy metal stress in plants. For example, Zhang, et al. (2018) determined that HMSSI was better at distinguishing heavy metal stress than when CI\_RE and PSRI were used alone. However, in that study, the HMSSI index was not able to quantify heavy metal stress during all plant growth stages and the results of this study further support this observation. Like all indices evaluated in this study, the HMSSI index was only able to quantify Cd stress during one time point, and it was only able to detect Cd stress in basil when images were taken from the top view (Table 2.5). Consequently, we conclude that it does not appear to be a particularly strong index for detecting Cd stress in leafy green crops.

## **2.6 Conclusions**

Leafy green crops like basil and kale can accumulate levels of Cd in edible plant tissues that greatly exceed health standards set by the FAO (Baldantoni et al., 2016) while displaying few symptoms of plant stress. The biochar amendment evaluated in this trial appeared to have subtle effects on reducing Cd stress in basil, but did not reduce Cd uptake in either crop indicating that amending soils with a rate 3% (v/v) will not be effective in protecting human health in contaminated soils. Several of the plant physiological indices evaluated in this trial appear to have some merit for detecting the subtle effects of Cd and soil amendments like biochar. In particular, NDVI holds promise for detecting Cd stress in basil and CI\_RE may be able to detect Cd stress in both basil and kale. PSRI and ARI both appear to have potential for quantifying the effects of Cd and biochar amendments in both crops. Being able to take images from the top rather than the side view of plants, appears to be the most effective way of capturing subtle changes in plant physiological responses in these two leafy green crops. These studies were conducted in a highly controlled environment, where we were able to remove any variation caused by other potential

stress factors such as nutrient, water, or pest pressure. They were also conducted using only one, artificially constructed soil. Separating effects of heavy metal stress from other environmental factors using remote sensing technology and without knowing prior information has been difficult in the past (Zhang et al., 2018). However, a distinct feature of heavy metal stress is that it is generally persistent, whereas other stress factors, such as nitrogen stress, are transient and only last for a short specific period of time (Zhang et al. 2018), so it is theoretically possible that these indices will be of value in the field. Consequently, additional studies will be needed to confirm that the indices identified in this can quantify the effects of Cd and biochar amendments in the presence of other plant stress factors, and future studies should be conducted in the field.

## **CHAPTER 3. IDENTIFICATION OF MATHEMATICAL MODELS THAT CAN QUANTIFY CADMIUM CONCENTRATION IN TWO LEAFY GREEN CROPS**

### **3.1 Abstract**

Cadmium (Cd) is a heavy metal that can get into the human body through the food chain, endangering human health. Phytotoxicity due to Cd stress is difficult to detect with the naked eye, making it difficult to determine whether plants are at risk, or if efforts to soil reduce bioavailability and uptake into edible plant tissues are working. New hyperspectral imaging (HSI) technologies developed for use in high-throughput plant phenotyping is a potent tool for monitoring environmental stress and predicting the nutritional status in plants. These imaging techniques have also proved valuable as non-invasive and autonomous approaches to detect biotic and abiotic stress during the early stages of plant stress symptoms. Previous studies have indicated the HSI can be used to detect Cd stress in plants like leafy greens that are at risk for Cd uptake, but it is unclear if this technology could be used to predict Cd concentrations in plant tissues. Consequently, this study was conducted to determine if mathematical models could be developed using HSI images to predict Cd concentrations in two leafy greens (kale and basil) amended or not with biochar, and subject to one of four soil Cd concentrations (0, 5, 10, and 15 ppm). The three models investigated were: principal components analysis (PCA), partial least squares (PLS) and artificial neural networks (ANN). Results of these studies indicate that the PCA and PLS models overfit (modeling error) the data despite efforts to transform the data in ways to predict more subtle signs of plant stress. In contrast, the ANN model was able to predict whether leafy greens had levels of Cd that were above or below critical thresholds suggested by the Food and Agriculture Organization (FAO), indicating that HSI could be used to predict Cd stress with this model.

### **3.2 Introduction**

Heavy metals are among the most dangerous pollutants in the environment due to their high levels of biological toxicity and persistence over time. Heavy metals can destroy the normal functioning of soils, disrupting critical processes like nutrient cycling, and cause severe stress in crops, impeding their growth and productivity (Wang et al., 2018). Cadmium (Cd) is one of the

most phytotoxic heavy metals, because of its potential to inhibit critical physiological processes such as photosynthesis, respiration, and the absorption, transport, and assimilation of mineral nutrients and water (Wang et al., 2018). Cadmium interferes with gene and protein expression, induces or inhibits enzymatic activity, increase the accumulation of reactive oxygen species and cause lipid peroxidation (He et al., 2015). Cadmium can also affect nitrogen (N) metabolism by inhibiting nitrate absorption and reducing the activity of enzymes involved in the nitrate assimilation pathway (He et al., 2015). Moreover, if heavy metals accumulate in plant tissues, they can also enter the food chain and harm human health.

It has been estimated that plant consumption of Cd contributes from 70 to more than 90% of the total intake of Cd by humans (Baldantoni et al., 2016). Leafy vegetables like spinach, lettuce, basil and kale are particularly problematic because they considered to be high Cd accumulators (Baldantoni et al., 2016). This is because these plants tend to have relatively high potential for Cd absorption and translocation, resulting in the accumulation of this toxic element in aboveground edible plant tissues. Leafy vegetables are an important component of human diet, and therefore they are an important source of Cd intake for people. The Food and Agriculture Organization (FAO) has determined that concentrations higher than  $0.28 \text{ mg kg}^{-1}$  of Cd in leafy vegetables can pose a serious threat to human health (Gu et al., 2015; FAO, 2018; Baldantoni et al., 2016). One of the biggest challenges in dealing with the threat of Cd in edible plant tissues like leafy greens is that toxicity symptoms can be difficult to detect. For example, while severe symptoms of Cd toxicity in plants can include growth retardation, chlorosis, necrosis, blackening of the roots and even death (He et al 2015), other plants can show no symptoms at all. Many Brassica species like kale appear to be particularly tolerant of Cd, displaying no visible symptoms of toxicity or significant reductions in plant biomass, even in the presence of high levels of soil Cd (Haghighi et al., 2016).

Currently, the most common and effective way for quantifying Cd in plant tissues is by post-harvest wet chemical extraction and analysis with atomic adsorption (AA) or inductively coupled plasma (ICP). However, these techniques are time consuming and expensive. An alternative approach that could be faster and less expensive is the application of sensing-based reflectance to quantify Cd concentration in plant foliage. With this approach, electromagnetic radiation is reflected from a target and quantified using images (Wang et al., 2018). The digital images developed using these detection technologies have proved useful in quantifying plant populations

in the field, and in various phenotyping applications such as quantifying plant canopy closure (Neilson et al., 2015 ; Zhou et al., 2019). Most of these images are currently obtained using standard red, green and blue (RGB) images, though researchers are now taking advantage of hyperspectral imaging, which can capture much more information (Lowe et al., 2017). Hyperspectral imaging is now being used to detect petroleum hydrocarbons and landmines (Wang et al., 2018). However, the potential for hyperspectral imaging to detect and estimate heavy metal concentrations in plants is in its infancy.

To understand how to develop imaging tools that can effectively detect and quantify heavy metals like Cd in plants, researchers must first understand how these images are obtained. For example, RGB images are composed of blue light (approximately 475 nm), green light (520nm) and red light (650 nm) (Lowe et al., 2017). These three primary colors are part of the visible (VIS) light spectrum and are also visible to humans. RGB imaging systems rely on cones, which are photoreceptor cells that respond to different wavelengths. These cones are particularly sensitive to blue, green, and red bands, and capture these wavelengths in the brain. According to the light wavelengths emitted, the cone stimulation will be either strong or weak, and the colors perceived will have a high or low saturation or intensity. RGB technology integrates the intensity and saturation of these three primary colors in a pixel. A leaf image contains hundreds of pixels. These 3-band-multispectral cameras based on RGB spectra have often been used to detect earlier stages of plant stress symptoms because they can track growth patterns, the greenness of plants, and parameters such as leaf area (Lowe et al., 2017).

Hyperspectral imaging (HSI) captures hundreds of contiguous narrow band wavelengths in a broader spectral range within each pixel (Lowe et al., 2017). These cameras can capture wavelengths in the VIS, as well as near-infrared regions (NIR) that go from 400 to 1400 nm. By capturing these wavelengths it is possible to detect changes in the leaf pigmentation (400–700 nm), and mesophyll cell structure (700–1300 nm) (Lowe et al., 2017). Between the VIS and the NIR region, is the red edge position. The red edge refers to the 680–750 nm wavelength range, where the spectral reflectance of vegetation increases sharply due to the absorption of red radiation by chlorophyll, and the strong reflection of infrared (IR) radiation (Wang et al., 2018). Capturing the red edge position is expected to be crucial for capturing stress caused by heavy metals (Wang et al., 2018). This is because there is a strong correlation between the red edge and leaf chlorophyll, the amplitude of the red border and carotenoid concentrations, and the peak area of the red border

and the indices related the leaf area and quality of fresh leaves (Wang et al., 2018). The red edge position moves to a longer wavelength when the vegetation is vigorous and has abundant chlorophyll, but it changes to a shorter wavelength if the plant is under stress (Wang et al. 2018).

Because of the high resolution, HSI cameras can generate a lot of environment noise; therefore, the spectral reflectance data generated by these cameras requires optimization. Optimizing reflectance data with spectral pretreatments will be particularly important for detecting heavy metal stress, since the responses are often very subtle and not always distinctive enough for detection (Zhao et al., 2018). Commonly used optimization methods include enhancement transformation, curve smoothing, first and second derivatives, continuous curve elimination, and wave-based noise elimination (Zhao et al., 2018). For example, the red edge position is often determined using the wavelength of the first maximum derivative in the VNIR range (Wang et al. 2017). Developing models that can detect heavy metals will also require identification of spectral bands that are most sensitive to heavy metal content. Hyperspectral data has a high degree of data redundancy among available bands, which makes identification of the most significant bands difficult. So far, the spectral bands that are the most effective for detecting Cd in plants are still unclear.

Some have suggested that the spectra of the first derivative that are found in the yellow border (IR 700–900 nm) are correlated with Cd, while others have suggested bands in the NIR is important (Zhao et al. 2018). In a recent study by Wang et al. (2018) that investigated spectral bands correlated with Cd contamination in *Brassica rapa chinensis* leaves, the researchers found that bands within the 690–1300 nm were sensitive, though 554, 631 and 557 nm appeared to the bands that were most highly correlated (Wang et al. 2018). Identification of the most sensitive and useful spectral bands can be performed by calculating the correlation between the heavy metal content measured *in situ* using traditional wet chemistry-based techniques with the raw spectral reflectance obtained during plant imaging (Wang et al. 2018). There are many approaches available to select wavelengths for plant stress analysis. The most typical model used to detect typical wavelengths is based on partial least squares regression (PLSR) analysis. This approach has been widely used to identify critical wavelengths among HSI images to quantify the quality of food and agricultural products (Olabe et al., 2005). However, it is necessary to explore other models that can model complex nonlinear relationship between inputs and outputs without

assuming specific distribution or independence of input variable (M.-I. B. Lin et al., 2012) like ANN, to determine if they could be more accurate.

Three alternative models that could identify wavelengths that can detect Cd stress include principal component analysis (PCA), partial least squares (PLS), artificial neural network (ANN) as they are standard parametric statistical prediction approaches. Principal component analysis (PCA) is a statistical tool that simplifies a model with many dimensions into a model with a linear combination and rotation of the initial vectors so that in the new coordinate the first few dimensions accounts for the majority of the variance in the dependent variable (Jolliffe et al., 2015; M.-I. B. Lin et al., 2012; Olabe et al., 2005, ). PCA models are valuable because they reduce complexity. However, the main problem with PCA is that it can be difficult to validate results since there is no response variable contrasting them. Partial least squares (PLS) models combine characteristics of principal component analysis (PCA) and multiple regression, and generalize the data (Wang et al., 2018; Zhou et al., 2019). They are useful because they can avoid the collinearity of simple regression models. PLS models transform predictor variables into uncorrelated orthogonal components, and perform least-squares regressions on these components rather than the original data, thus decreasing the size of the predictor variable space and preserving the same information of the original variables. Artificial neural networks (ANN) are mathematical models of artificial intelligence that are based on the biological behavior of neurons and structure of the human brain (Olabe et al., 2005) The elements of ANN models behave similar to the basic structure of biological neurons and present a series of characteristics of the human brain.

ANN models are particularly interesting because they can be used to generate knowledge about a particular component based on previous studies, or exercises and experience. For example, the program is given a set of inputs and runs until it can produce consistent outputs. Because ANN can generalize automatically, it can provide precise answers even though the input variables cause some small disturbances or noises. ANN models can separately consider qualities of an object that does not have common or relative aspects. The neural system is made up of the following elements: neurons, connection pattern, memory and learning dynamics, and environment. The analog unit of a biological neuron is the PE processor component, which has several inputs and combines them by means of a basic addition operation. These are modified by a transfer operation, and return output values that go directly to the output of the processor element. The PE outputs connect to the inputs of other artificial neurons for the efficiency of the neuronal synapse. The ANN objective

is not only in the processor component, but also in the way in which they are connected. In general, PE is connected in the form of layers, and a traditional neural network is a sequence of connections between very close consecutive layers. The underlying neural architecture consists of an input layer or input buffer where the network data is presented and an output layer or output buffer where the network response to the input data is presented. The rest of the layers are called hidden layers. ANN can be viewed as universal model-free approximators, that can represent any nonlinear function with sufficient accuracy by seeking the proper combination of several sigmoid functions (Shi et. al 2014). It has the advantage of self-learning, robustness and self-organization in modeling (Shi et al., 2014). ANN is rarely used for the quantitative analysis of plant stress and properties (Shi et al., 2014) However, the ANN model has been used the most in image processing, and is appropriate for large multivariate data sets Shi et al., 2014) and therefore has strong potential to identify Cd stress. Also, for this model we are able to preselect some wavelengths beforehand and use those for modeling. As we use fewer layers in the ANN we reduce dimensions.

Not all Cd that is in soil is bioavailable and therefore can be taken up by plants. Consequently, identifying strategies that can reduce the bioavailability of this heavy metal in soil can reduce human health risks. For example, previous studies have demonstrated that soil amendments like biochar can reduce Cd uptake in leafy greens (Gomez and Hoagland, unpublished). However, to identify effective strategies for soil Cd immobilization, researchers need rapid and cost-effective approaches to detect Cd stress and estimate plant uptake.

Hyperspectral imaging has potential to more rapidly and cost-effectively quantify Cd in plant tissues. However, making this a reality will require identification individual wavelengths that can most accurately detect Cd stress. This requires the use of statistical and mathematical tools to evaluate the large amount of data generating when using HSI. Therefore, the goal of this study is to identify mathematical models that can most accurately detect Cd stress and estimate Cd concentrations in the edible tissues using kale and basil as two representatives but distinctly different leafy green crops.

### **3.3 Materials and methods**

#### **3.3.1 Experimental design and plant elemental concentrations at harvest**

The data used to identify models that would best fit the data and predict Cd concentrations in this chapter was presented in chapter 2. Briefly, the data came from a plant growth experiment conducted at Purdue University's Controlled Environment Phenotyping Facility (CEPF) in West Lafayette Indiana, U.S. using a growth media derived of equal parts by volume of sand, soil and BM8 potting media mix. Half of these pots received a locally sourced biochar at a rate of 3% v/v, and the other half were left untreated, before pots with both soil treatments were amended with CdCl<sub>2</sub> to obtain concentrations of 0, 5, 10, and 15 ppm total growth media Cd. After a brief incubation period to allow Cd to adsorb onto soil particles, the pots were planted with either basil or kale and placed in a carefully controlled growth chamber to facilitate plant growth. Each plant species X addition of biochar X Cd level was replicated four times. After approximately three months, the plants were destructively harvested to quantify total aboveground biomass and concentrations of C, N, Cd and other elements.

#### **3.3.2 Elemental concentrations in aboveground plant biomass**

Total carbon (C) and nitrogen (N) in kale and basil aboveground biomass was quantified after subjecting 0.5 g samples of dry biomass to combustion at 840 C (LECO, CE Elantech, Lakewood, NJ, USA). Concentrations of total Cd and zinc (Zn) in plant tissues were determined using ICP-OES (Shimadzu ICPE-9820 and location) following digestion using a Mars 6 (CEM, Charlotte NC, USA) with Xpress vessels. Briefly, 0.5 g samples were placed in 10 ml HNO<sub>3</sub> and subject to a temperature of 200 °C, a pressure of 800 psi, and a power of 900-1050 watts.

#### **3.3.3 Manual measurements and images collected during the experiment**

Plants spectral data were collected using a hyperspectral camera (400-998 nm) using a Middleton Spectral Vision Camera (Middleton Spectral Vision, Middleton, Wisconsin, USA). This HSI camera is a VIS + IR camera capable of sweeping from the 400 to 998 nm wavelengths, with a spectral resolution of 473 bands. Generated data generated swas collected by a program called Smarter AgTM (Purdue AgIT, West Lafayette, IN, USA). Each plant was imaged using two different views (top and side view), and the system automatically collected images at different

heights depending on the size of each plant. Leaf area was estimated using images collected from the top and side view of basil and kale plants during the experiment at three time points.

### 3.3.4 Spectral pre-processing

To increase the quality of the hyperspectral images it was necessary to decrease environmental and physiological factors (noise), since this affect the signals in the hyperspectral data (Zhao et al 2018). For this to happen, data pre-processing is required. To obtain the spectral spectrum, the data was normalized using white and dark reference images using the following formula:

$$R = \frac{R_0 - R_D}{R_W - R_D}$$

In this formula  $R_0$  is the raw reflectance data,  $R_W$  is the white reference data,  $R_D$  is the dark reference data and  $R$  is the relative reflectance (Zhao et al., 2018)

In addition, to promote the signal-to-noise ratio, data at the beginning and end of the spectral bands were removed (Zhao et al., 2018). The regions representing 457 to 980 nm were retained, and the spectral bands representing before 457 and greater than 980 nm, were removed.

### 3.3.5 Standard normal variation (SNV)

The average reflectance spectra were first transformed to absorbance  $\log_{10} \left( \frac{1}{reflectance} \right)$  using methods described in Zhao et al. (2018). We transformed reflectance data into absorbance data because absorbance data is related to plant activity, while reflection is more related to the plant surface. Then, to eliminate the undesirable impacts such as random noise, light scattering, and baseline shifts (Zhao et al., 2018), three spectral preprocessing techniques were performed: detrending, Savitzky–Golay first derivative and standard normal variate (SNV). SNV was applied to remove the scatter effect. Savitzky–Golay first derivative was used to remove the baseline shift and amplify small spectral features (Zhao et al. 2018). Detrending (the elimination of trending data) was used to eliminate the effects of baseline shifts and curvilinearity (Zhao et al. 2018). Then the first derivative and the detrending with normalization (SNV) were combined. The combination of SNV + detrending is used to remove curvilinearity and absorbance offsets from NIR spectra (Zhao et al., 2018) to show the essential information that we are looking for

### 3.3.6 Mathematical models

For the PCA and PLS models, 10- fold cross-validation was performed to unify the data effectiveness. Normalization of data and the first derivative were also applied, and the models were performed using the spectral responses of kale and basil leaves, and the total amount of Cd found in the leaves. For the ANN model, correlation coefficients and the root mean square error (RMSE) were transformed to better predict Cd stress in kale and basil plants. For all models, data collected from the kale and basil plants were combined due to the low number of samples before performing regression and classification activities. An illustration taken from Shi et al., (2014) (see Figure 3.1), illustrates the technical approaches used in this study to estimate Cd in plants using hyperspectral images.

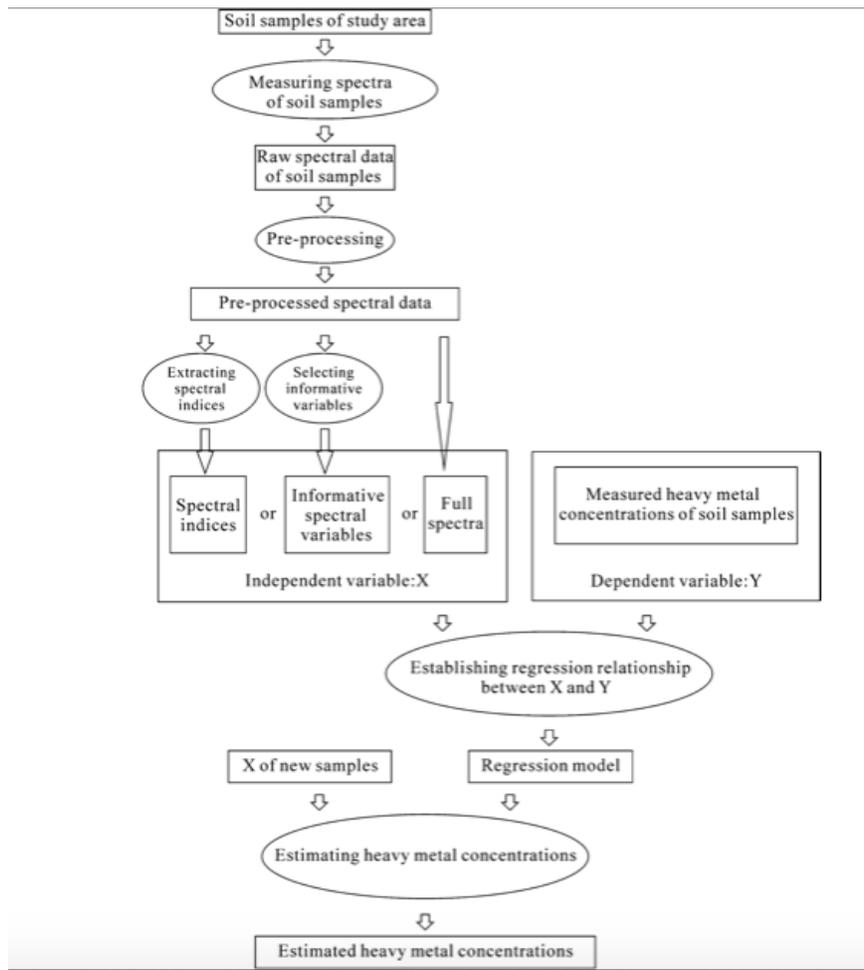


Figure 3.1 Diagram obtained from Shi et al. (2014) illustrating how hyperspectral imaged can be used to conduct experiments to develop models for estimated heavy metal concentrations

### 3.3.7 Statistical analyses

Statistical differences in plant leaf area and cadmium and nitrogen concentrations were determined using the statistical software R and various R- packages. ANOVA was initially used to quantify differences in the model to identify correlations and significance between treatments was determined using the least significance difference (LSD) test. MATLAB 2019 was also used to perform the PCR and PCA mathematical models, and Python 3.8 was used to perform the ANN model.

## 3.4 Results and Discussion

### 3.4.1 Cadmium and nitrogen concentrations in basil and kale biomass at harvest quantified with traditional chemical extraction and analyses

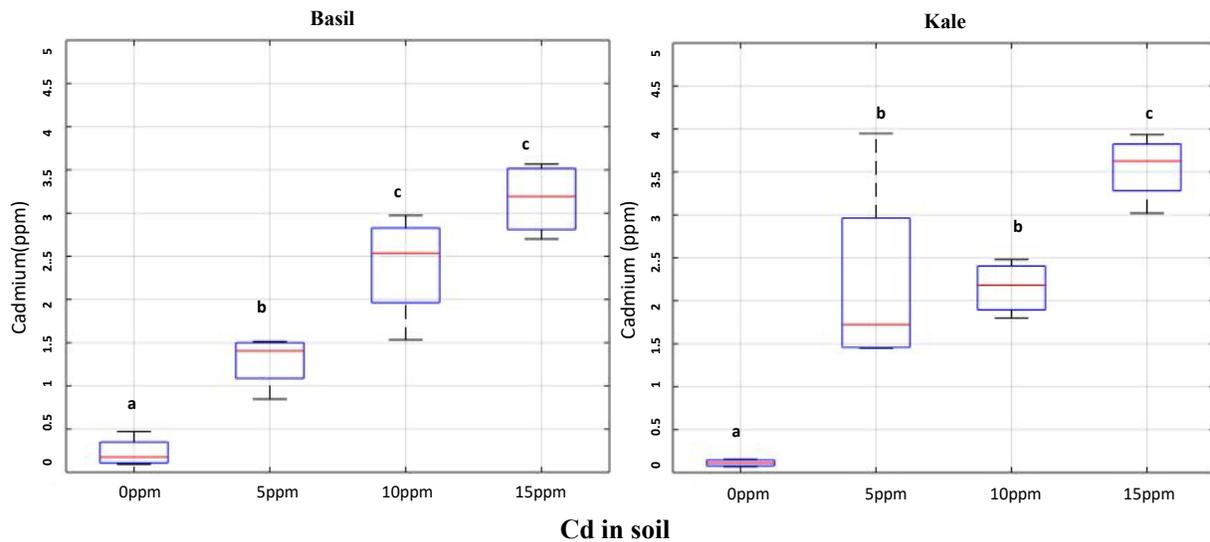
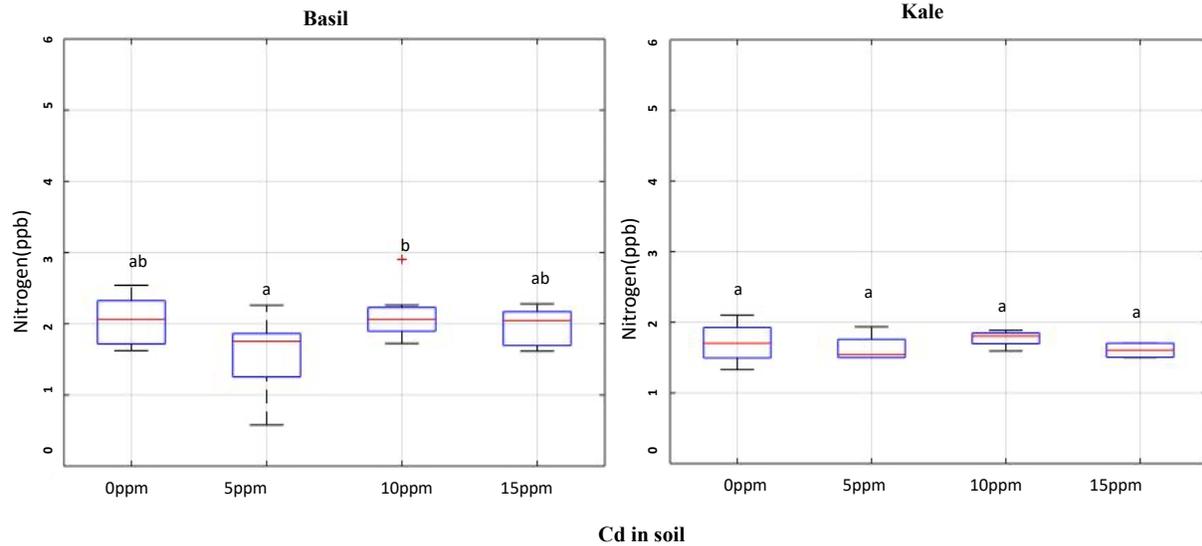


Figure 3.2 Boxplots indicating concentrations of cadmium and nitrogen obtained using ICP-OES in the aboveground biomass of basil and kale plants subject to four soil Cd concentrations. Different letters indicate significant differences at the  $P < 0.05$ .

Figure 3.2 continued



Results of the ICP-OES analyses demonstrate that there was a positive correlation between soil Cd concentrations and plant accumulation of Cd in the aboveground tissues of both kale and basil plants (Figure 3.2), demonstrating that these are good plant species for use in developing the models described in this chapter. Because of the strong negative relationship that Cd can have on plant nitrogen pathways, which could affect the potential for HSI to detect Cd stress, the concentration of N in basil and kale leaves was also determined. However, in this study, there were no significant differences in N concentrations among the soil Cd concentrations for either plant, indicating that Cd soil levels in this study may not have been high enough to generate severe symptoms of Cd toxicity symptoms (Figure 3.2). Nevertheless, results of the previous chapter demonstrated that it is possible to use HSI imaging to detect subtle differences in Cd stress in these two leafy green crops, even though this does not translate to changes in overall N.

### 3.4.2 Changes in basil and kale leaf area in response to biochar and soil Cd levels

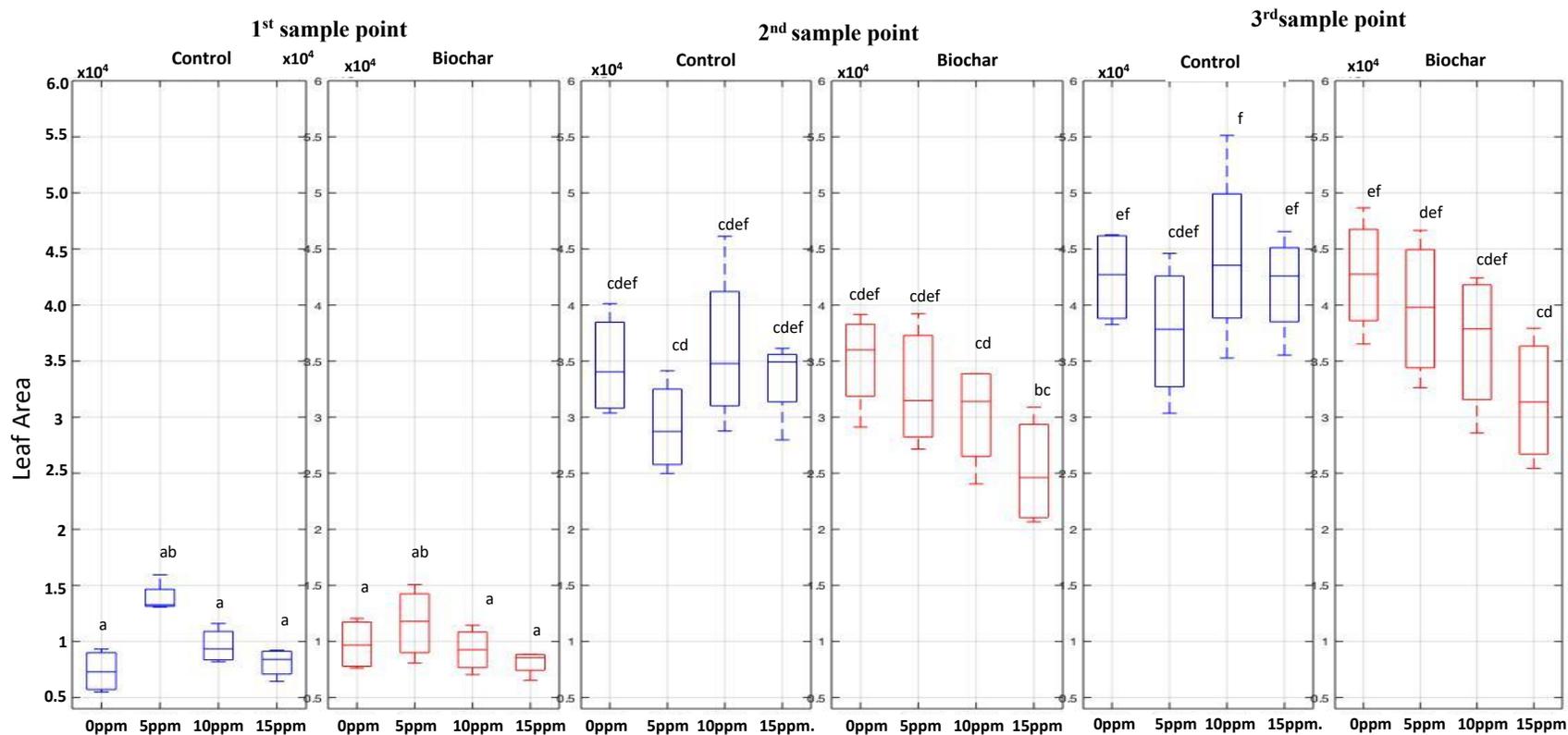
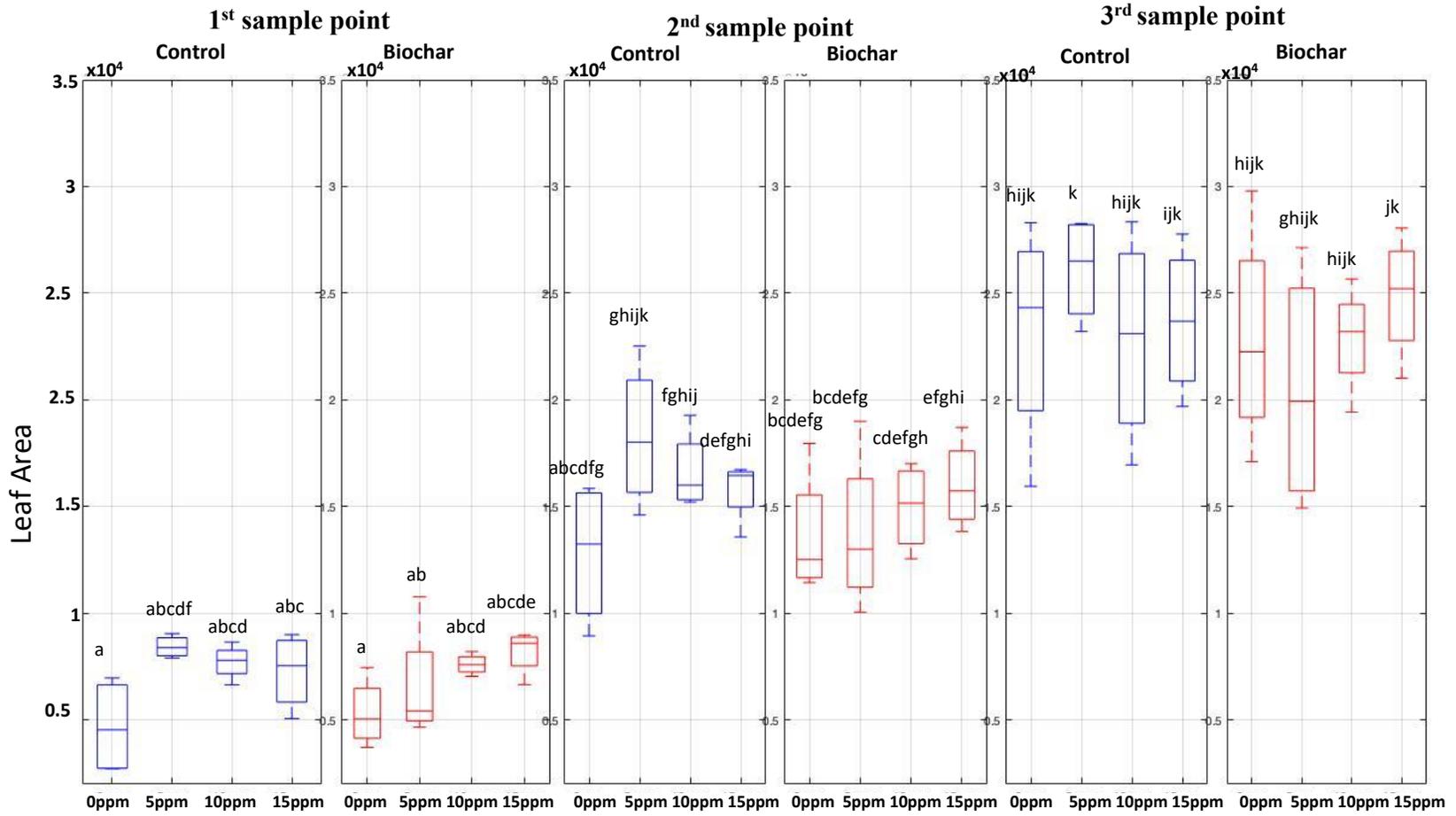


Figure 3.3 Changes in basil (top panel) and kale (bottom panel) leaf area in plants amended or not with biochar and subject to four soil Cd concentrations at three time points. Different letters indicate significant differences at the  $P < 0.05$

Figure 3.3 continued



Leaf area is an important indicator of plant stress that could be used in models to estimate Cd concentrations. Consequently, leaf area was estimated using images collected from the top and side view of basil and kale plants during the experiment at three time points. Only the results of the analyses using images take from the side view are presented here, as they better illustrate the results (Figure 3.3).

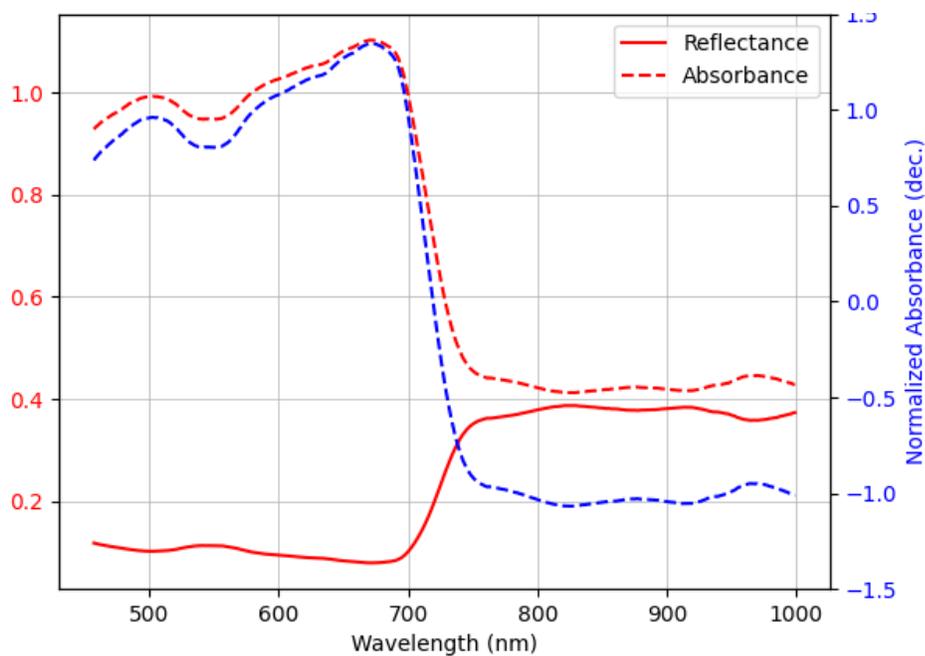
There were few differences in leaf area among basil plants in response to addition of biochar and the four soil Cd concentrations, indicating that leaf area is not a strong predictor of Cd stress in this crop. However, leaf area was significantly lower in the plants that were amended with biochar and subject to 15 ppm Cd during the third time point, than those that did not receive the biochar amendment addition but were subject to the same rate of soil Cd. As described in the previous chapter, plants have been demonstrated to increase growth in response to low levels of plant stress in a phenomenon known as hormesis, which has been theorized to be an evolutionary adaption (Agathokleous et al., 2019). Consequently, it is possible that the biochar amendment was reducing Cd stress at the highest soil Cd level.

Among kale plants, there were no significant differences in response to the biochar addition among soil Cd treatments within individual time points; however, there were significant differences in growth rates over time, indicating the kale leaf area is more sensitive to Cd stress than basil (Figure 3.3). Specifically, between the first and second time points, in the absence of the biochar addition, leaf area was significant greater in the second time point when plants were subject to 5, 10 and 15 ppm Cd, but there was no difference when plants were amended with biochar. These results indicate that the soil amendment could have reduced early stimulations in plant growth caused by Cd stress, which could also be related to hormesis.

### **3.4.3 Spectral pre-processing to develop models to predict Cd concentrations in plant foliage**

Previous studies investigating differences in leaf reflectance induced by heavy metals were not significant, and the authors concluded that plant responses to heavy metals are often so subtle that they are not sensitive enough to develop models for estimating heavy metal concentrations in plant tissues (Wang et al., 2018; Zhou et al. 2018). One of the reasons that these studies failed to detect a significant difference, is environmental-induced noise generated during the imaging process, which is generally higher than the subtle features associated with plant heavy stress (Zhou

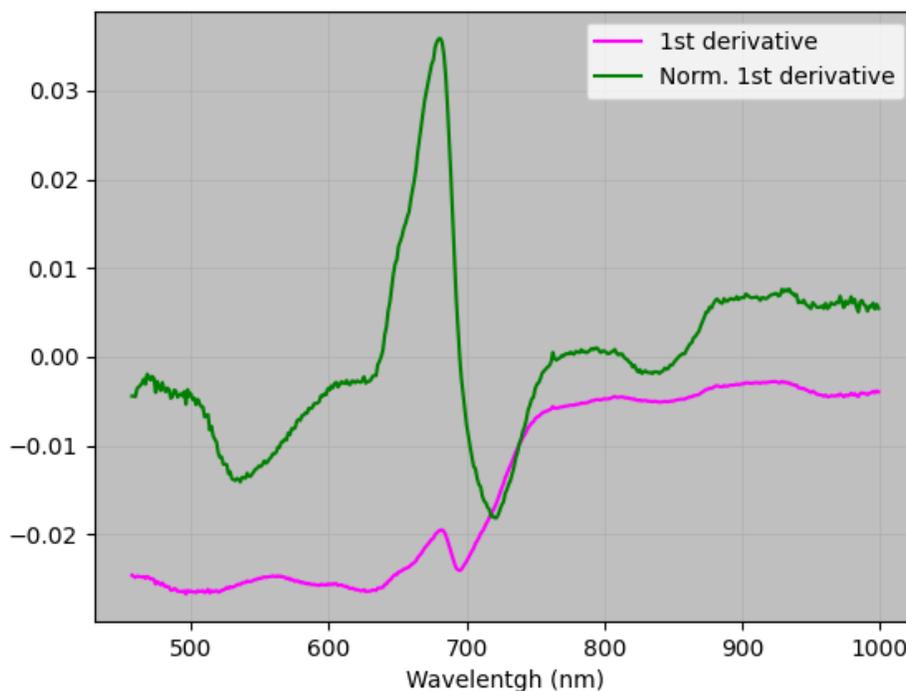
et al. 2018). For example, during HSI, factors such as random noise, light scattering and baseline shifts can make it difficult to interpret data (Zhou et al., 2019). Therefore, it is necessary to remove noise and amplify stress response information in the original reflectance data (Zhou et al., 2019). Consequently, three spectral preprocessing techniques were used in this study to reduce the effects of environmental noise, and the reflectance data was transformed into absorbance data to increase the accuracy of the images. In addition, because Cd concentrations in both basil and kale plants were positively correlated with soil Cd levels, the data from the two plant species were combined to increase accuracy in the model.



*Figure 3.4 Raw reflectance and absorbance data and normalized absorbance generated using a random sample of basil and kale plants subjected to soil Cd stress*

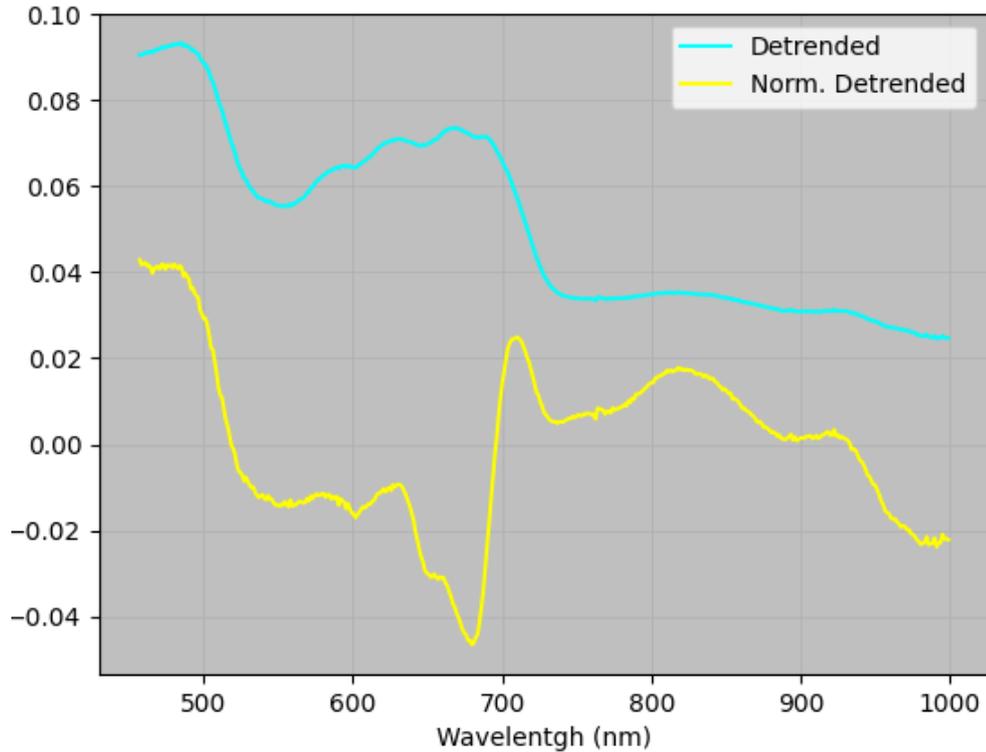
Results of our efforts to reduce environmental noise and transform the data into information useful for developing models to predict Cd concentrations in basil and kale are shown in Figure 3.4. The blue line illustrating a normalized correlation coefficient in the absorbance data where we see a decrease in the reflectance compared to the raw absorbance data, especially with respect to wavelengths in the infrared region. This indicates that we were successful in our attempt to eliminate much of the environmental noise in this data set. To further eliminate background

interference, resolve overlapping spectra, and minimize the noise and baseline drift associated with the raw spectral data, we combined the SNV for the first derivative of absorption spectra. The first derivative can increase spectrum correlations but improve baseline shifts and amplify small spectral features, and it can also reduce the signal-to-noise ratio, especially in the 800-900 nm regions (Zhao et al. 2018). Using data from a random plant (Figure 5) it is possible to visualize differences in the absorption index between the first derivative and the normalization of the derivative data. Both of these curves had a peak when shifting from visible to the infra-red range. Cd does generally not cause plants to have unique absorption bands in the NIR region, though long-term exposure of plants to Cd may affect leaves chlorophyll synthesis, which could be detected (Zhao et al., 2013; Zhou et al., 2018). Most of the spectral features associated with plant chlorophyll are concentrated in the spectral region associated with the red-edge, such as the inflection point of the red edge (Zhao et al., 2018). Some of the transformations we performed appeared to decrease the noise to signal ratio, allowing us to better characterize subtle differences at the red edge. However, we could still see some interference in the normalized first derivative.



*Figure 3.5 Spectral data generated using a random basil or kale plant subject to soil Cd stress. The top panel represents the first derivative and first derivative + SNV absorbance data, and the right panel represents the detrending and detrending + SNV absorbance*

Figure 3.5 continued



Combining SNV and detrending of HSI data can be used to remove curvilinearity and absorbance offsets from the near-infrared region (NIR) to better characterize subtle differences at the red edge. For example, in Figure 3.5, it is possible to visualize differences between the absorption of detrending + normalized detrending data next to the detrending absorbance data alone. In this figure, the normalized first derivative (First derivative + SNV) and detrended data had a higher absorbance than the 1<sup>st</sup> derivative and normalized detrended data, indicating that this technique can better clarify responses. However, it was still possible to detect noise in the normalized detrending data (detrending + SNV), and in the normalized first derivative (first derivative + SNV), indicating that it is still difficult to detect subtle differences.

### 3.4.4 Developing mathematical models to estimate Cd concentration in leafy green crops

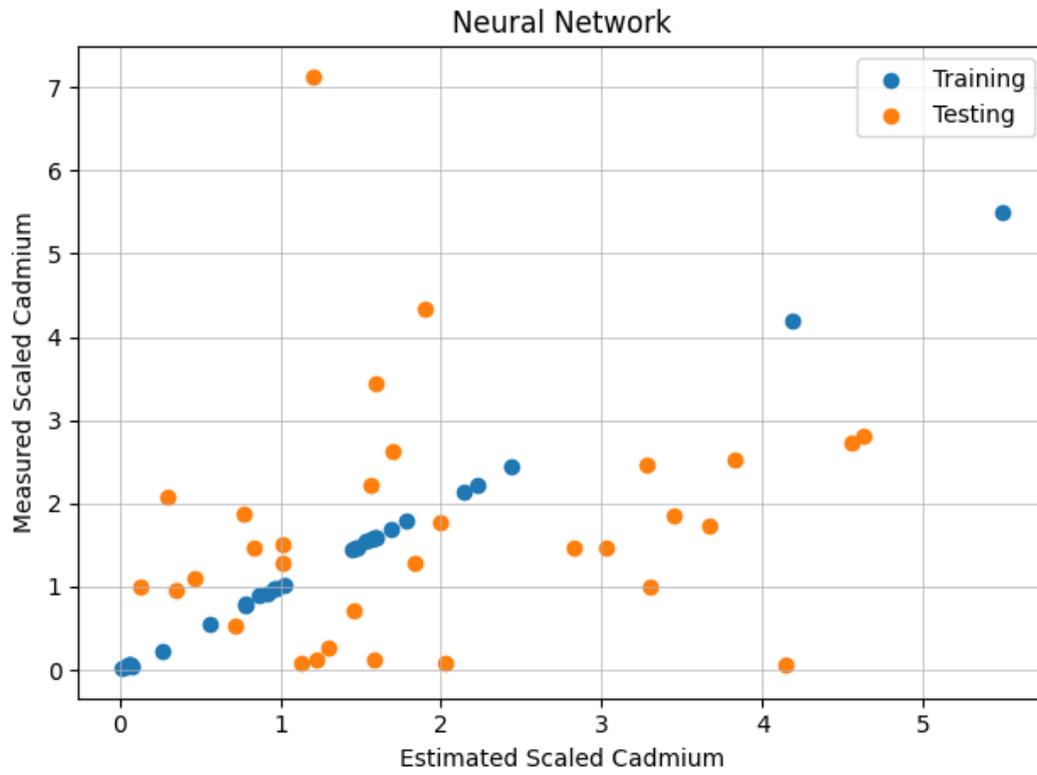


Figure 3.6 ANN regression model generated using the detrending + SNV transformed data of plants subject to soil Cd stress

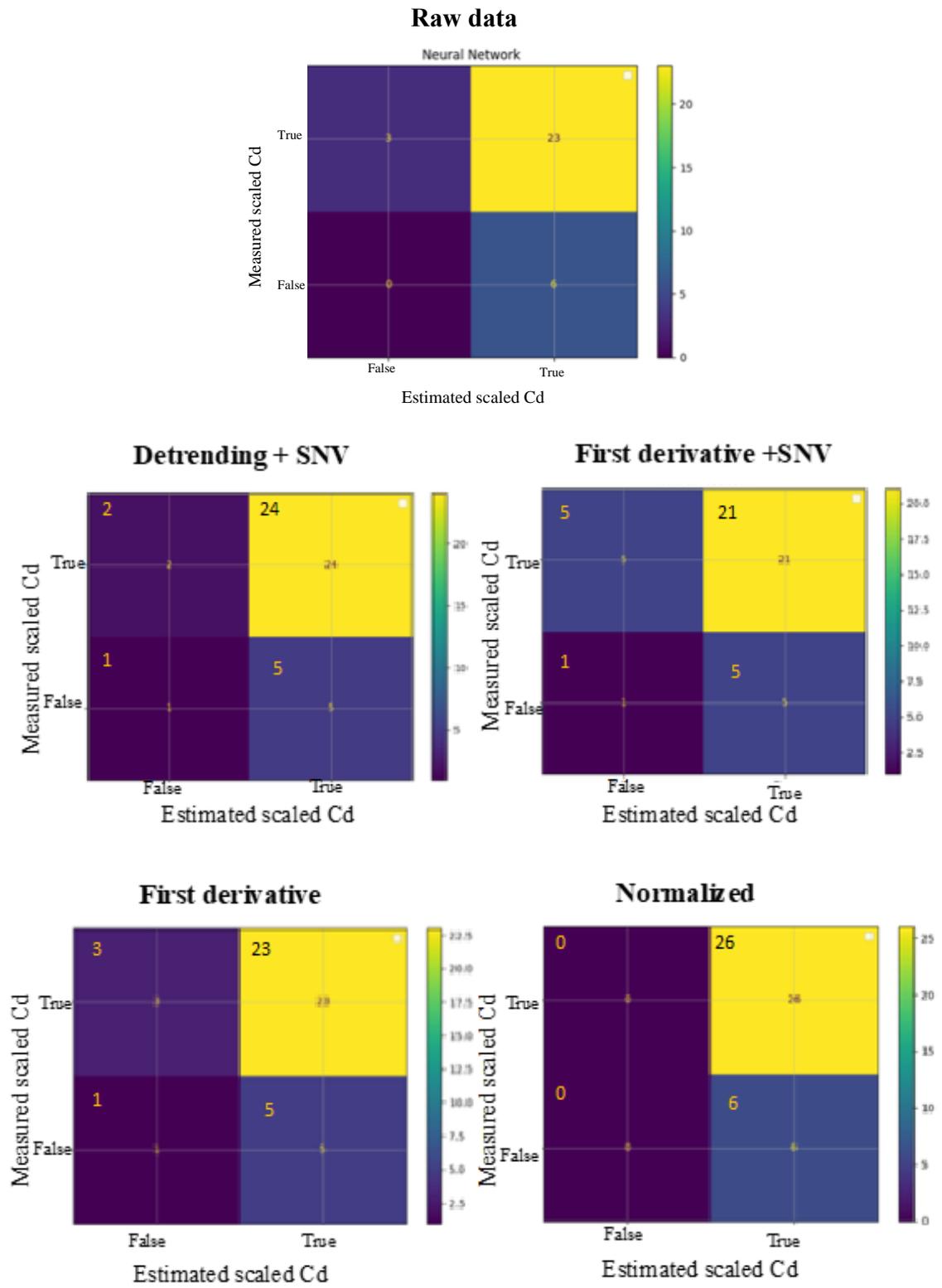
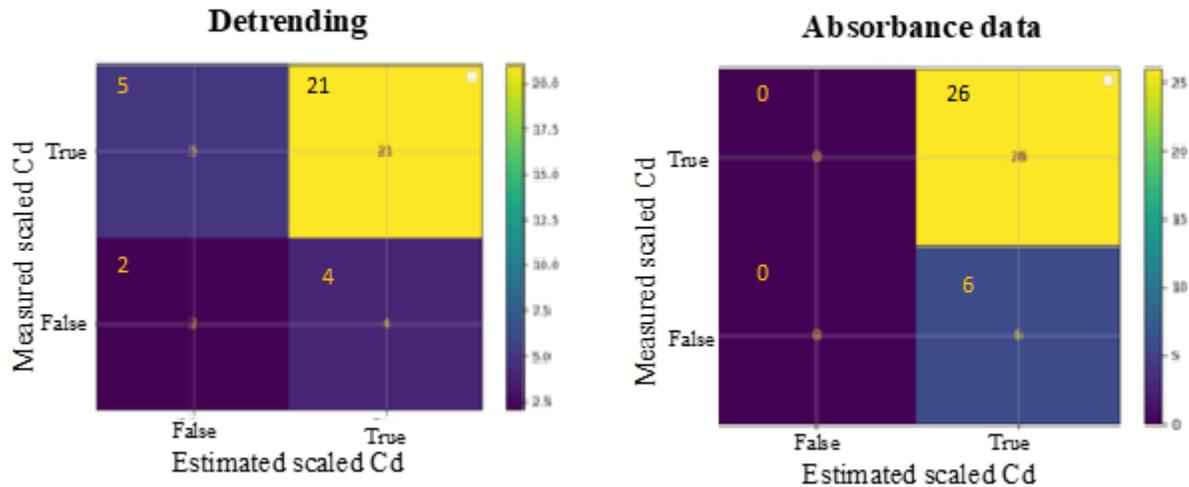


Figure 3.7 Confusion matrices generated using the ANN model and different data transformations using leafy greens subject to Cd stress.

Figure 3.7 continued



Once all the data was transformed to improve accuracy, it was loaded into the three models evaluated here and further adjusted to attempt to improve the fit of the models. For example, for the ANN model, we conducted a regression for every data transformation, but were unable to identify any data set that could fit the model and predict Cd stress accurately (Figure 3.6) With machine learning, we were able to develop learning algorithms that could predict outcome values for previously unseen data. However, since only a finite number of samples were used, the evaluation of these samples can be sensitive to sampling error. As a result, measurements of prediction error on the current data may not provide much information about predictive ability on new data.

Consequently, we decided to perform a confusion matrix using an ANN classification to see which one could best predict Cd concentrations in the plants (Figure 3.7). We used this approach because we wanted to determine if the model could classify plants that were considered health and unhealthy with respect to Cd concentrations. The threshold used in this analysis was taken from the FAO (Baldantoni et al., 2016), which states that concentrations of C greater than 0.28 mg per kg of fresh weight is considered dangerous in leafy greens. The data were considered “true” if Cd concentrations in plant tissues exceeded maximum levels ( $0.28\text{mg kg}^{-1}$ ), and “false if it did not exceed this value. The model was divided into “tested” data, that which was obtained directly from the plant growth experiments, and “trained” data, which was generated by the program. All analyses were performed using random data from each data set. This model was able

to We classified the plants that had higher or lower Cd concentration than the threshold in the leafy green crops evaluated in this study (Table 3.1).

*Table 3.1 Statistics table demonstrating the power of these models to detect Cd concentrations in basil and kale aboveground foliage using the transformed data.*

	Precision Training	Recall Training	F1-score Training	Precision Testing	Recall Testing	F1-score Testing
Det and normalize	1	1	1	0.828	0.923	0.873
First Derivative and normalized	1	1	1	0.808	0.808	0.808
Detrend	1	1	1	0.84	0.808	0.824
First derivative absorbance	1	1	1	0.821	0.885	0.852
Normalized absorbance	0.893	0.962	0.926	0.8125	1	0.897
Absorbance	0.812	1	0.900	0.812	1	0.897
Reflectance	1	1	1	0.793	0.885	0.834

After performing the ANN model using the classification method and different data transformation approaches, we were able to quantify differences in their predictive value (Table 3.1). This generated an F1 score, precision and recall information from the training and testing data. Recall is the ratio of correctly predicted positive observations to all observations in an actual class while precision is the ratio of correctly predicted positive observations to the total predicted positive observations (Ping Shung, 2018). F1 is the weighted average of Precision and Recall. Therefore, precision is a good measure to determine if the costs of false positive are high. Recall is was the model metric we used to select our best model when there was a high cost associated with false negatives, since recall actually calculates how many of the actual positives the model was able to capture through labeling the data as positive (True Positive) (Ping Shung, 2018).. For example, the detrending + normalized (SNV) data set had a recall and F1 score of 100% for the training data and for 92 % the testing data, indicating that the model was very precise. The combination of absorbance and normalized absorbance (SNV) also had a high recall and F1 score for both the training and testing data sets. The detrending + SNV absorbance approach was able to predict 24 values that exceeded the Cd threshold out of 26, and 1 of the false values out of 6 that did not exceed maximum recommended limit. In contrast, the normalized absorbance (SNV) and

raw absorbance were able to predict 26 out of 26 true values, but only 0 out of 6 false values. Most of the approaches had high scores, which indicates that ANN is a good method to detect when leafy greens exceed the maximum level of Cd suggested by the FAO. For example, all of them had a recall testing value that was higher than 80%, indicating that they can classify more than 80% correctly. Consequently, we conclude that this classification method, derived using an artificial network, is a good model for detecting leafy greens that are considered safe for human health. Further testing to validate this hypothesis is underway using synthetic data to balance the range of Cd values

The PCA model was conducted using 8 principal components, and the PLS model was conducted using both 4 and 3 principal components, and both the side and top view images were evaluated in both models to determine which view could provide a better fit for the data. Figure 3.8 illustrates the relationship between the measured and predicted Cd concentrations when images were taken from the top view, and Figure 3.9 represents the relationship when images were collected from the side view. Previous studies have demonstrated that PLSR can be effectively used to estimate heavy metals including Cd (Wang et al., 2018). For example, Zhou et al. (2019) constructed a PLSR model using rice leaves spectrums that was able to estimate Cd concentrations in brown rice. Similarly, Wang et al. (2018) used a PLSR model to overcome the limitation of other ordinary models in addressing collinearity among explanatory variables, and this model had greater accuracy in estimating heavy metal content. However, the pace of improvement was related to the type of metals. In Cd, the model had an  $R^2$  of 0.69 to 0.72 in rice. Wang et al. (2018) found that a PLSR model produced better results than an ANN model, because the ANN model was unable to identify the unique contribution of selected variables to the dependent variable. In our study, data was divided randomly into training (70%) and validation (30%) sets. When we created different PCA and PLSR models with these different sets, the sampled data fit well, where the  $R^2$  of the training data was high, but this all created an overfitting model. Also the root mean square deviation (RSME) was high in all the models created. This means that the generalization error was high, indicating that the model may not provide much information about the predictive ability towards new data. Moreover, the top view in kale and side view for both models overfitted, indicating that these models were accurate, however, the validation data was not and therefore needing further analysis in order for these models to have value in detecting Cd stress in these crops.

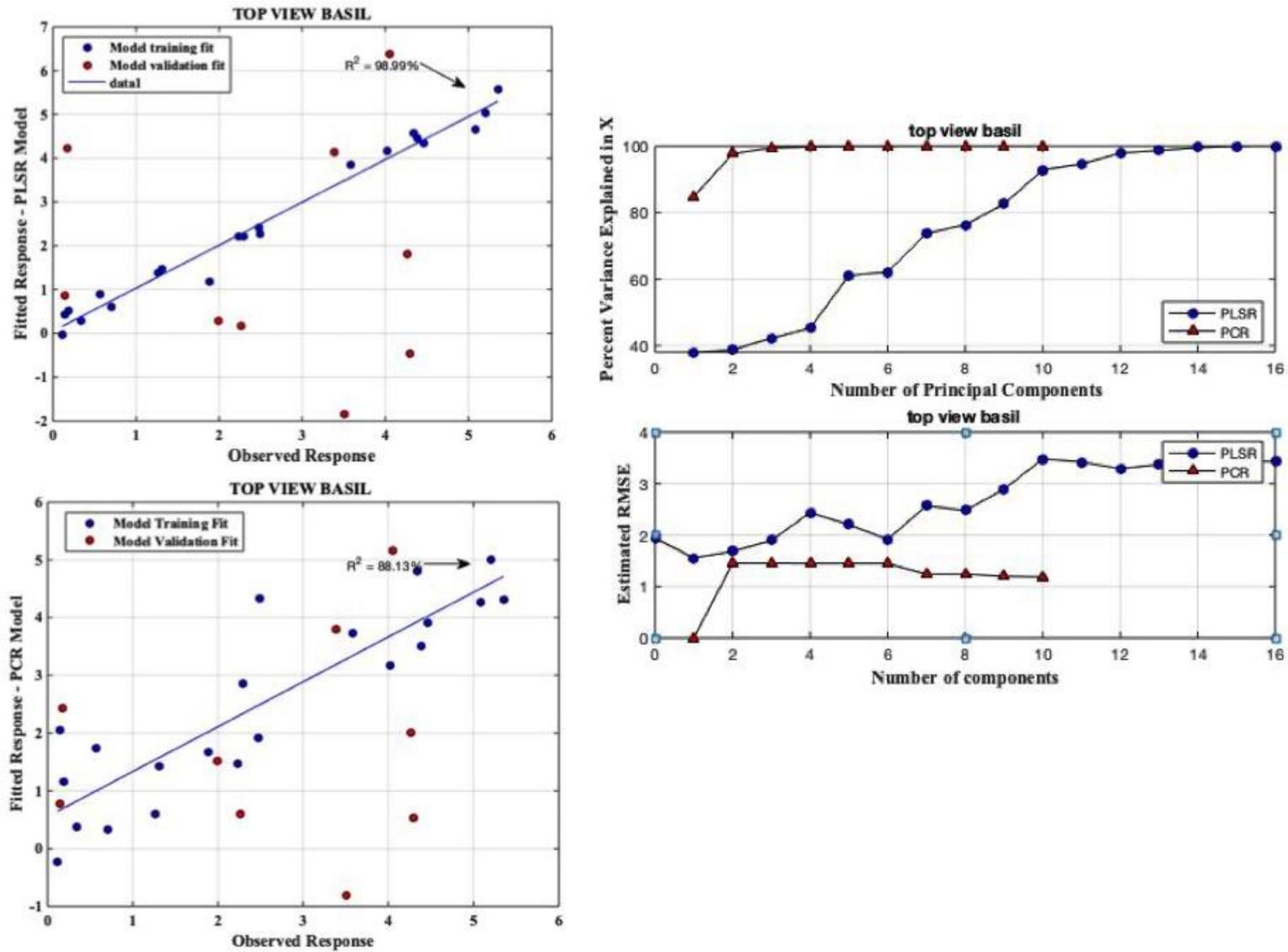
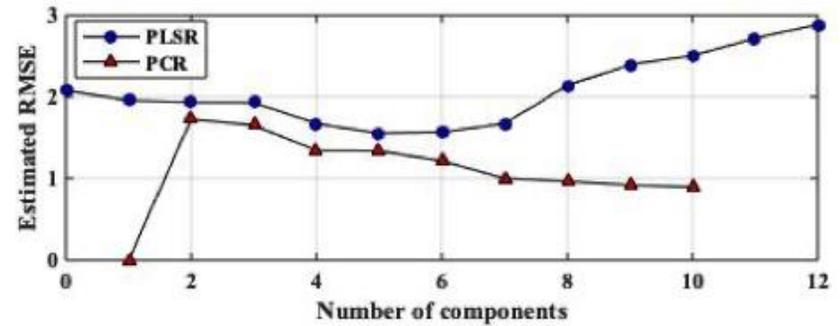
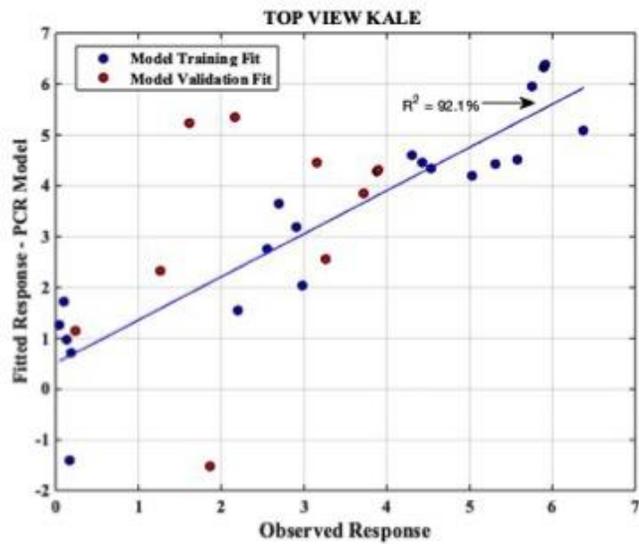
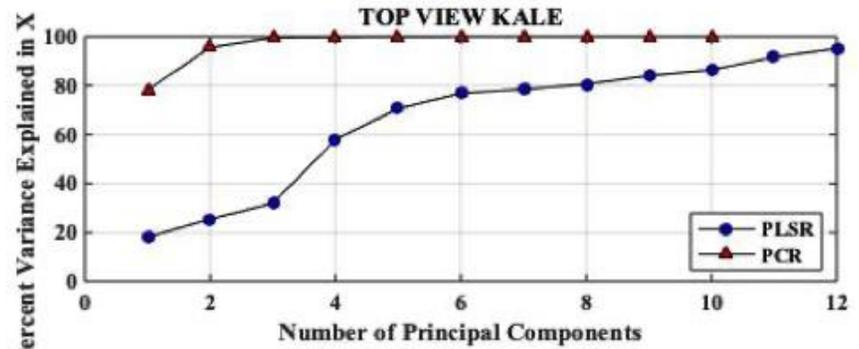
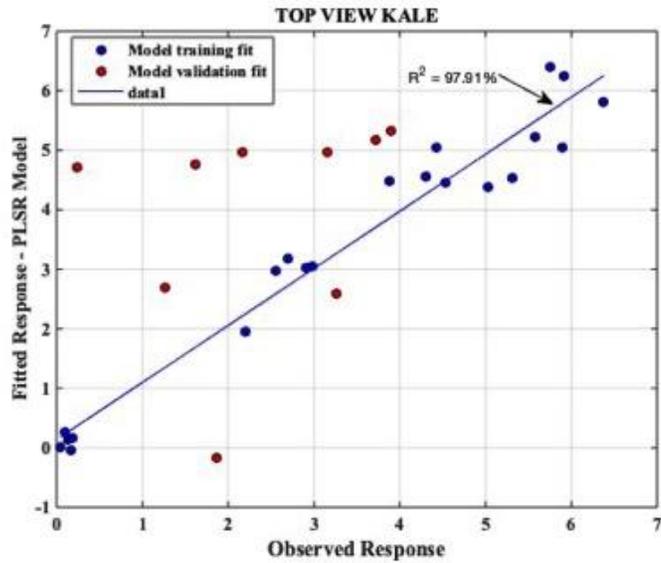


Figure 3.8 8 PCA and PLS models developed using basil and kale plants subject to soil Cd stress and images collected top view of plants.

Figure 3.8 continued



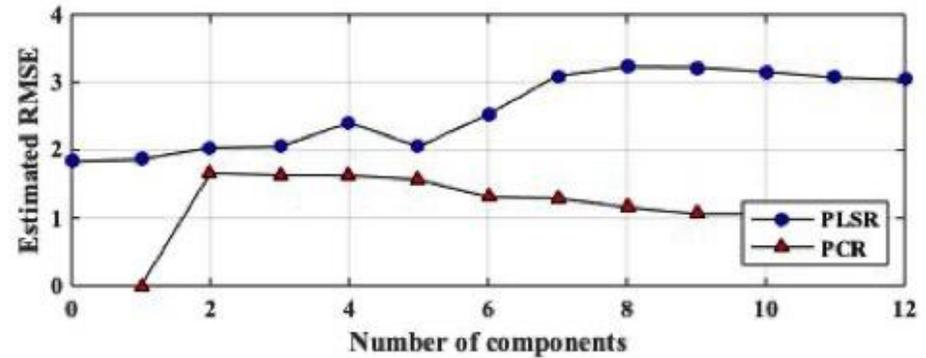
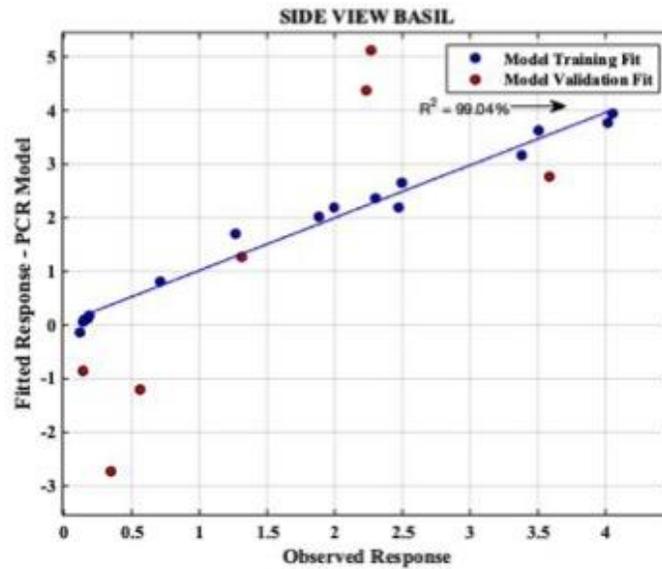
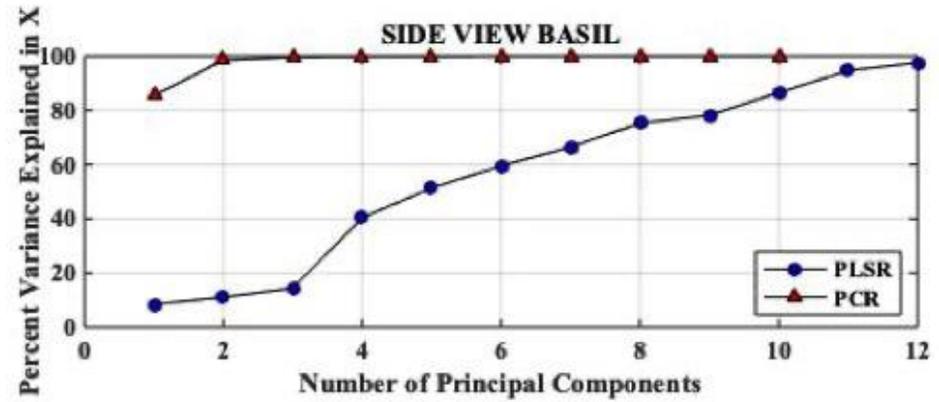
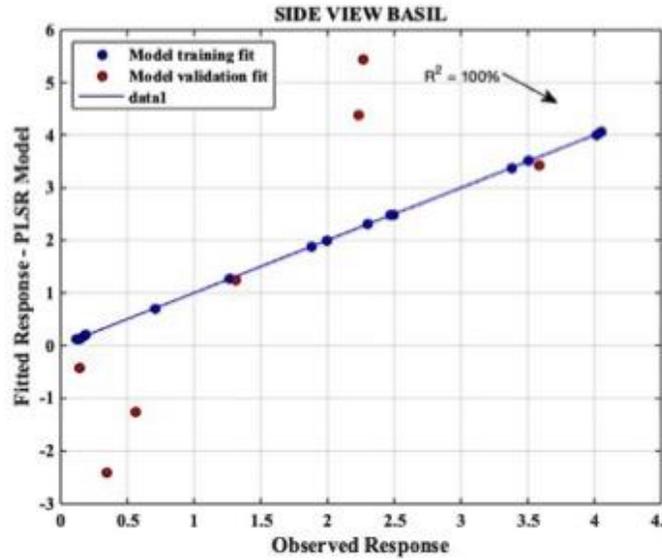
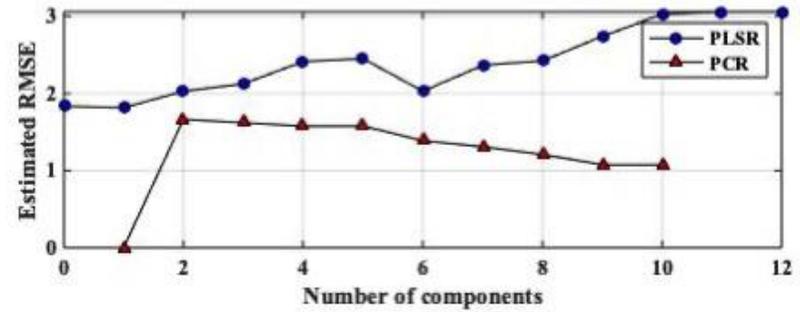
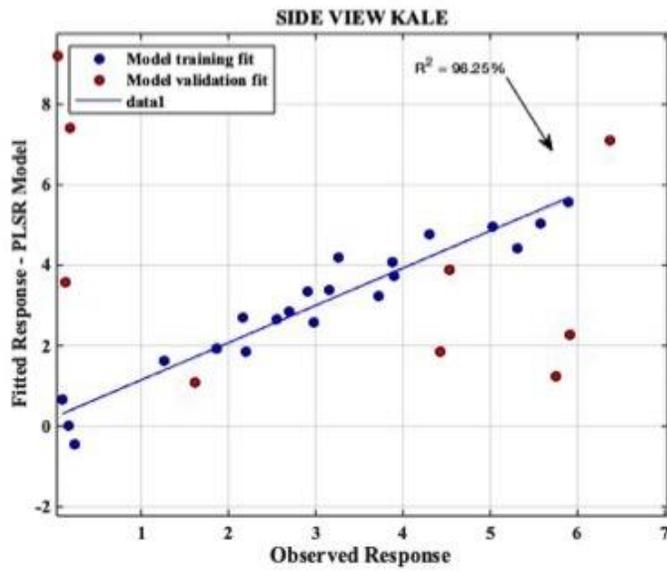
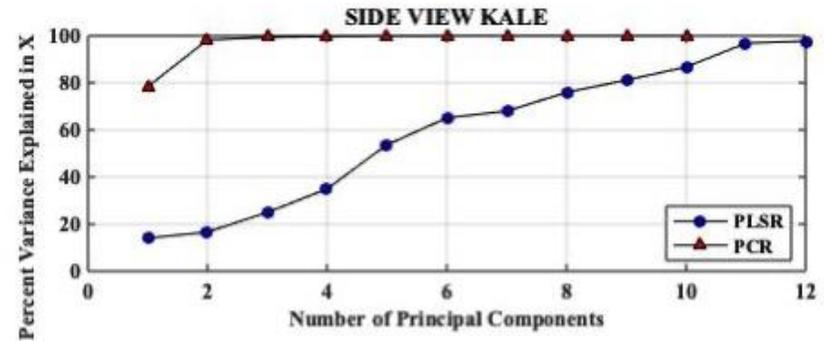
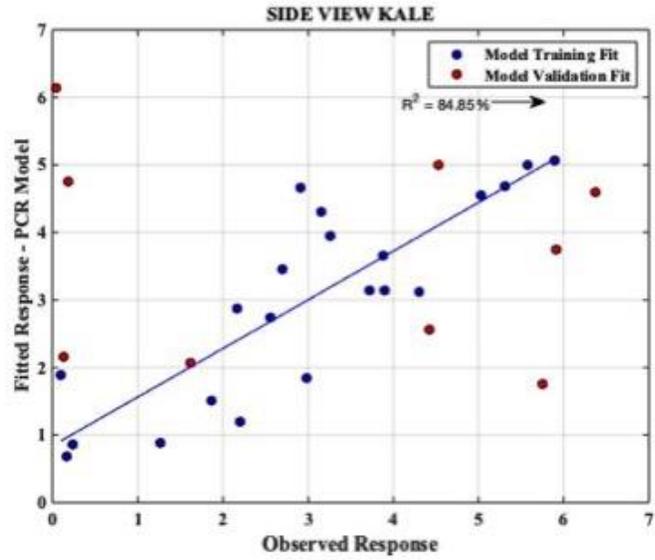


Figure 3.9 PCA and PLS models developed using basil and kale plants subject to soil Cd stress and images collected side view of plants

Figure 3.9 continued



### **3.5 Conclusions**

The PCA and PLS models constructed using hyperspectral images of kale and basil were able to detect plant stress due to Cd toxicity, but they were not effective in predicting Cd concentrations in the leaves of these plants. Cross-validation statistic procedures and normalization of the data increased the accuracy of these models to a certain degree and helped reduce the signal-to-noise ratio data, but they still over fit the data. Consequently, we conclude that PCA and PLS were not useful for predicting Cd concentrations in the foliage of these two leafy green crops. In contrast, the ANN model, which often has greater accuracy in predicting data because it combines several sigmoid functions, was able to predict, at an accuracy of more 80%, whether these leafy greens had Cd levels that were considered safe and unsafe by the FAO. This indicates that this model could be incredibly helpful in identifying food lots that could be contaminated, and could be useful in research studies aimed at reducing soil bioavailability and uptake of Cd. However, since these trials were conducted under highly controlled conditions using a small number of samples, further studies conducted under real world conditions and with larger sample sizes are needed to validate this assertion and further improve the predictive capacity of this model

## CHAPTER 4. CONCLUSION

Cadmium is a toxic element that is detrimental to the health of soil, plants and humans. New methods for detecting and quantifying Cd bioavailability and uptake into edible plant tissues is critical to resolving these challenges. Not all plants display visible symptoms of Cd stress, and current methods for quantifying Cd concentrations in plant tissues involve a long process of destruction and wet chemistry, which is expensive and time-consuming.

The research conducted as part of this thesis project demonstrate that hyperspectral imaging has potential for use in detecting subtle changes in plant stress caused by Cd, as well as potential Cd stabilization processes in soil (see chapter 2). Fourteen different indices generated using the HSI system were evaluated to determine which might be most helpful for detecting Cd stress in kale and basil. Results of this study confirmed that not all of them were useful, though NDVI was able to detect Cd stress in basil and CI\_RE was able to detect Cd stress both basil and kale, indicating that these might be the most helpful indices for use in these crops. HMSSI, developed explicitly for heavy metal detection, was able to detect Cd stress in basil, but only when images were taken from the top view. Results for the red edge indices: MSRE\_RE and NDVI\_RE, were also only able to detect stress when images were taken from the side view. Therefore, we conclude that the location where images are taken is critical for the accuracy of hyperspectral vegetative indices in detecting Cd stress in these two leafy green crops. The top view images had better results than images collected from the side view, so future efforts to optimize this methodology should rely on images collected from the top view. In addition, all of the indices only detected differences in plant stress responses when plants were young, indicating that future efforts to optimize this technology should be conducted early.

The second goal of the study described in this thesis was to determine whether mathematical models generated using the HSI images could be used to predict Cd concentrations in the leaves of kale and basil at harvest. Three mathematical models were evaluated, though only one, the Artificial Neural Network (ANN) model, was successful in predicting Cd concentrations. Interestingly, the model appeared to be able to predict whether kale and basil leaves contained Cd that exceeded values considered by the FAO (0.28 mg/kg of fresh weight) as safe for human consumption. This indicates that HSI technologies could be very useful in future studies aimed at Cd remediation efforts.

Effective ways to immobilize Cd in soil is critical for preventing the uptake of this element on contaminated soils. In many places of the world, moving production to non-contaminated areas is not practical. Previous studies have provided evidence that some types of biochar can immobilize soil Cd and prevent uptake in crops. In these studies, we investigated whether a local biochar amendment derived from hardwood trees could reduce Cd uptake in kale and basil leafy greens, but results indicated that the rate we investigated was not effective. However, we did observe subtle differences in basil plants with respect to the disappearance of a stress response phenomenon, which we attribute to a phenomenon known as hormesis, indicating that the amendment could have some benefits, perhaps if higher rates were applied. In addition, we did observe that some vegetative indices (PSRI and ARI) were able to distinguish differences in plants amended with biochar and plants that were not. Interestingly, PSRI and ARI both quantify differences in secondary metabolites, which are often produced by plants under stress and can be helpful in mitigating challenges associated with stress, such as making plants more resistant to pathogens. Previous studies have demonstrated that biochar can make plants more resistant to pathogens, so it would be interesting to test these relationships in future studies.

While the studies described in this thesis were conducted using leafy green crops, the long-term goal of this project is to address the issue of excess Cd in cacao cropping systems in Colombia. Cacao plants are much larger and perennial, which makes them difficult to work with in a controlled chamber like what was used in this study. However, given that we observed most plant responses occurred when plants were young, we suspect that it may be possible to conduct these assays on young cacao plants in future research projects. A small experiment was conducted to see if the system could work on cacao plants and preliminary results indicate that it will be possible to generate HSI images on these plants (Figure 4.1) In the meantime, it may be possible to grow sensitive leafy green crops on cacao farms as indicator crops to detect high Cd levels, or determine whether remediation strategies are working. In addition, it may be possible to use hyperaccumulator leafy greens like kale, as part of a phytoremediation strategy on cacao farms. In either case, after further optimization of the system, results of these studies indicate that HSI could be used in these approaches. However, this would also require extensive testing in the field to validate this approach, as in this study, we were able to eliminate other potential plant stresses caused by factors such as nutrients, water or pest pressure, which could also affect the indices evaluated in this study.

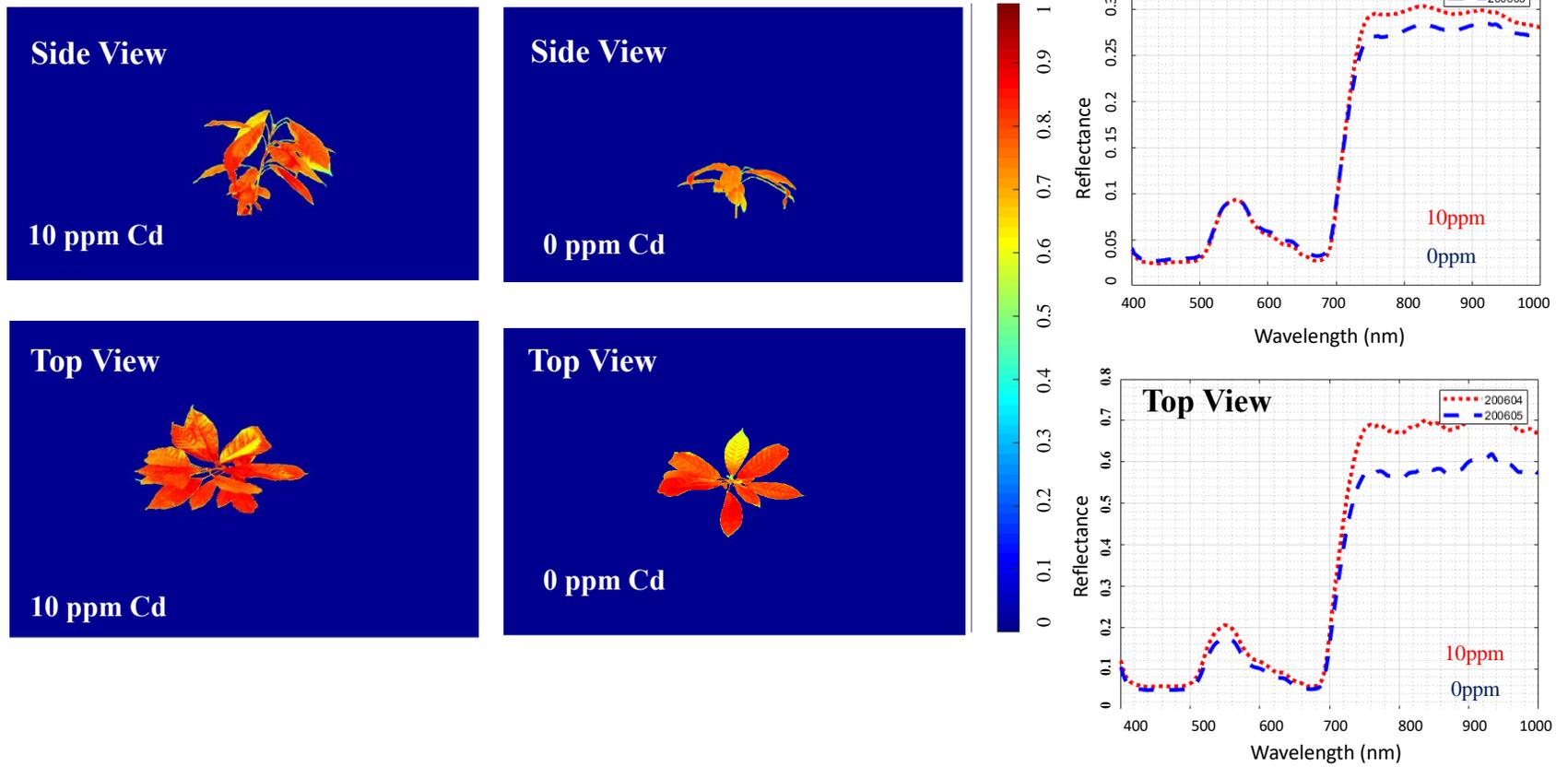


Figure 4.1 NDVI color maps and reflectance graphs of cacao plants grown in soil subject to 0 or 10 ppm soil Cd.

## BIBLIOGRAPHY

- Agathokleous, E., Kitao, M., Calabrese, E.J., 2019. Hormesis: A Compelling Platform for Sophisticated Plant Science. *Trends in Plant Science* 24, 318–327. <https://doi.org/10.1016/j.tplants.2019.01.004>
- Aprotosoiaie, A.C., Luca, S.V., Miron, A., 2016. Flavor Chemistry of Cocoa and Cocoa Products- An Overview: Flavor chemistry of cocoa .... *COMPREHENSIVE REVIEWS IN FOOD SCIENCE AND FOOD SAFETY* 15, 73–91. <https://doi.org/10.1111/1541-4337.12180>
- Arévalo-Gardini, E., Arévalo-Hernández, C.O., Baligar, V.C., He, Z.L., 2017. Heavy metal accumulation in leaves and beans of cacao (*Theobroma cacao* L.) in major cacao growing regions in Peru. *Science of The Total Environment* 605–606, 792–800. <https://doi.org/10.1016/j.scitotenv.2017.06.122>
- Ashourloo, D., Mobasheri, M., Huete, A., 2014. Evaluating the Effect of Different Wheat Rust Disease Symptoms on Vegetation Indices Using Hyperspectral Measurements. *Remote Sensing* 6, 5107–5123. <https://doi.org/10.3390/rs6065107>
- Baldantoni, D., Morra, L., Zaccardelli, M., Alfani, A., 2016. Cadmium accumulation in leaves of leafy vegetables. *Ecotoxicology and Environmental Safety* 123, 89–94. <https://doi.org/10.1016/j.ecoenv.2015.05.019>
- Bandaru, V., Daughtry, C., Codling, E., Hansen, D., White-Hansen, S., Green, C., 2016. Evaluating Leaf and Canopy Reflectance of Stressed Rice Plants to Monitor Arsenic Contamination. *IJERPH* 13, 606. <https://doi.org/10.3390/ijerph13060606>
- Benjamin, T., Lundy, M.M., Abbott, P.C., Burniske, G., Croft, M., Fenton, M.C., Kelly, C., Camayo, F.R., Wilcox, M., 2018. AN ANALYSIS OF THE SUPPLY CHAIN OF CACAO IN COLOMBIA. <https://doi.org/10.13140/RG.2.2.19395.04645>
- Bergsträsser, S., Fanourakis, D., Schmittgen, S., Cendrero-Mateo, M., Jansen, M., Scharr, H., Rascher, U., 2015. HyperART: non-invasive quantification of leaf traits using hyperspectral absorption-reflectance-transmittance imaging. *Plant Methods* 11, 1. <https://doi.org/10.1186/s13007-015-0043-0>
- Bot, A., Food and Agriculture Organization of the United Nations, Benites, J., Food and Agriculture Organization of the United Nations. Land and Water Development Division, 2005. Drought-resistant Soils: Optimization of Soil Moisture for Sustainable Plant Production : Proceedings of the Electronic Conference Organized by the FAO Land and Water Development Division, FAO land and water bulletin. Food and Agriculture Organization of the United Nations.
- Caligiani, A., Marseglia, A., Palla, G., 2016. Cocoa: Production, Chemistry, and Use, in: *Encyclopedia of Food and Health*. Elsevier, pp. 185–190. <https://doi.org/10.1016/B978-0-12-384947-2.00177-X>

- Chavez, E., He, Z.L., Stoffella, P.J., Mylavarapu, R.S., Li, Y.C., Moyano, B., Baligar, V.C., 2015. Concentration of cadmium in cacao beans and its relationship with soil cadmium in southern Ecuador. *Science of The Total Environment* 533, 205–214. <https://doi.org/10.1016/j.scitotenv.2015.06.106>
- de Souza, P.A., Moreira, L.F., Sarmiento, D.H.A., da Costa, F.B., 2018. Cacao— *Theobroma cacao*, in: *Exotic Fruits*. Elsevier, pp. 69–76. <https://doi.org/10.1016/B978-0-12-803138-4.00010-1>
- FEDECACAO, 2018. FEDERACION NACIONAL DE CACAOTERO INFORMACION CONSTITUCIONAL. FEDECACAO.
- Gharebaghi, A., Alborzi Haghighi, M., Arouiee, H., 2017. Effect of cadmium on seed germination and earlier basil ( *Ocimum Basilicum* L. and *Ocimum Basilicum* Var. *Purpurescens* ) seedling growth. *TJS* 15, 1–4. <https://doi.org/10.15547/tjs.2017.01.001>
- Gitelson, A.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, 271–282. <https://doi.org/10.1078/0176-1617-00887>
- Haghighi, M., Kafi, M., Pessarakli, M., Sheibanirad, A., Sharifinia, M.R., 2016. Using kale ( *Brassica oleracea* var. *acephala* ) as a phytoremediation plant species for lead (Pb) and cadmium (Cd) removal in saline soils. *Journal of Plant Nutrition* 39, 1460–1471. <https://doi.org/10.1080/01904167.2016.1161768>
- Hayyat, A., Javed, M., Rasheed, I., Ali, S., Shahid, M.J., Rizwan, M., Javed, M.T., Ali, Q., 2016. Role of Biochar in Remediating Heavy Metals in Soil, in: Ansari, A.A., Gill, S.S., Gill, R., Lanza, G.R., Newman, L. (Eds.), *Phytoremediation*. Springer International Publishing, Cham, pp. 421–437. [https://doi.org/10.1007/978-3-319-40148-5\\_14](https://doi.org/10.1007/978-3-319-40148-5_14)
- He, S., He, Z., Yang, X., Stoffella, P.J., Baligar, V.C., 2015. Soil Biogeochemistry, Plant Physiology, and Phytoremediation of Cadmium-Contaminated Soils, in: *Advances in Agronomy*. Elsevier, pp. 135–225. <https://doi.org/10.1016/bs.agron.2015.06.005>
- He, S., Yang, X., He, Z., Baligar, V.C., 2017. Morphological and Physiological Responses of Plants to Cadmium Toxicity: A Review. *Pedosphere* 27, 421–438. [https://doi.org/10.1016/S1002-0160\(17\)60339-4](https://doi.org/10.1016/S1002-0160(17)60339-4)
- Huang, W., Lu, J., Ye, H., Kong, W., Hugh Mortimer, A., Shi, Y., 1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China, 2. Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China, 3. Institute of Geographical Science, Hebei Academy of Sciences, Shijiazhuang 050021, China, 4. Hainan Provincial Key Laboratory of Earth Observation, Sanya 572029, Hainan, China, 5. Rutherford Appleton Laboratory, Harwell Science and Innovations Campus, Oxfordshire OX11 0QX, UK, 2018. Quantitative identification of crop disease and nitrogen-water stress in winter wheat using continuous wavelet analysis. *International*

- Journal of Agricultural and Biological Engineering 11, 145–152.  
<https://doi.org/10.25165/j.ijabe.20181102.3467>
- Huang, W., Yang, Q., Pu, R., Yang, S., 2014. Estimation of Nitrogen Vertical Distribution by Bi-Directional Canopy Reflectance in Winter Wheat. *Sensors* 14, 20347–20359.  
<https://doi.org/10.3390/s141120347>
- ICCO, 2020. ICCO Quarterly Bulletin of Cocoa Statistics (No. No. 2-Volume XLVI). International Cocoa Organization.
- ICCO, 2018. ICCO Quarterly Bulletin of Cocoa Statistic (No. No. 2-Volume XLIV). International Cocoa Organization.
- ICCO, 2013. growing cocoa. International Cocoa Organization (ICCO).
- ICCO, 2009. Boletín Trimestral de Estadísticas de Cacao (No. XXXV, N° 3). International Cocoa Organization (ICCO).
- Ismael, M.A., Elyamine, A.M., Moussa, M.G., Cai, M., Zhao, X., Hu, C., 2019. Cadmium in plants: uptake, toxicity, and its interactions with selenium fertilizers. *Metallomics* 11, 255–277. <https://doi.org/10.1039/C8MT00247A>
- Kumar, V., Marković, T., Emerald, M., Dey, A., 2016. Herbs: Composition and Dietary Importance, in: *Encyclopedia of Food and Health*. Elsevier, pp. 332–337.  
<https://doi.org/10.1016/B978-0-12-384947-2.00376-7>
- Lehmann, J., Joseph, stephen, n.d. Biochar for environmental managment.
- Lin, M.-I.B., Groves, W.A., Freivalds, A., Lee, E.G., Harper, M., 2012. Comparison of artificial neural network (ANN) and partial least squares (PLS) regression models for predicting respiratory ventilation: an exploratory study. *Eur J Appl Physiol* 112, 1603–1611.  
<https://doi.org/10.1007/s00421-011-2118-6>
- Lin, P., Qin, Q., Dong, H., Meng, Q., 2012. Hyperspectral vegetation indices for crop chlorophyll estimation: Assessment, modeling and validation, in: 2012 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2012 - 2012 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Munich, Germany, pp. 4841–4844.  
<https://doi.org/10.1109/IGARSS.2012.6352529>
- Loganathan, P., Vigneswaran, S., Kandasamy, J., Naidu, R., 2012. Cadmium Sorption and Desorption in Soils: A Review. *Critical Reviews in Environmental Science and Technology* 42, 489–533. <https://doi.org/10.1080/10643389.2010.520234>
- Lowe, A., Harrison, N., French, A.P., 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods* 13, 80. <https://doi.org/10.1186/s13007-017-0233-z>

- Mutanga, O., Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of Remote Sensing* 25, 3999–4014. <https://doi.org/10.1080/01431160310001654923>
- Neilson, E.H., Edwards, A.M., Blomstedt, C.K., Berger, B., Møller, B.L., Gleadow, R.M., 2015. Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C4 cereal crop plant to nitrogen and water deficiency over time. *Journal of Experimental Botany* 66, 1817–1832. <https://doi.org/10.1093/jxb/eru526>
- Ngala, T.J., 2015. EFFECT OF SHADE TREES ON COCOA YIELD IN SMALL-HOLDER COCOA (*Theobroma cacao*) AGROFORESTS IN TALBA, CENTRE CAMEROON.
- Peñuelas, J., Filella, I., Gamon, J.A., Field, C., 1997. Assessing photosynthetic radiation-use efficiency of emergent aquatic vegetation from spectral reflectance. *Aquatic Botany* 58, 307–315. [https://doi.org/10.1016/S0304-3770\(97\)00042-9](https://doi.org/10.1016/S0304-3770(97)00042-9)
- Ping Shung, K., 2018. Accuracy, Precision, Recall or F1?
- Ruf, F., zadi, H., n.d. Cocoa: From Deforestation to Reforestation.
- Sánchez-Pardo, B., Carpena, R.O., Zornoza, P., 2013. Cadmium in white lupin nodules: Impact on nitrogen and carbon metabolism. *Journal of Plant Physiology* 170, 265–271. <https://doi.org/10.1016/j.jplph.2012.10.001>
- Sharma, R.K., Agrawal, M., Agrawal, S.B., 2010. Physiological, Biochemical and Growth Responses of Lady's Finger (*Abelmoschus esculentus* L.) Plants as Affected by Cd Contaminated Soil. *Bull Environ Contam Toxicol* 84, 765–770. <https://doi.org/10.1007/s00128-010-0032-y>
- Shi, T., Chen, Y., Liu, Y., Wu, G., 2014. Visible and near-infrared reflectance spectroscopy—An alternative for monitoring soil contamination by heavy metals. *Journal of Hazardous Materials* 265, 166–176. <https://doi.org/10.1016/j.jhazmat.2013.11.059>
- Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* 81, 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X)
- Sridhar, B.B.M., Han, F.X., Diehl, S.V., Monts, D.L., Su, Y., 2007. Spectral reflectance and leaf internal structure changes of barley plants due to phytoextraction of zinc and cadmium. *International Journal of Remote Sensing* 28, 1041–1054. <https://doi.org/10.1080/01431160500075832>
- Sun, H., Feng, M., Xiao, L., Yang, W., Wang, C., Jia, X., Zhao, Y., Zhao, C., Muhammad, S.K., Li, D., 2019. Assessment of plant water status in winter wheat (*Triticum aestivum* L.) based on canopy spectral indices. *PLoS ONE* 14, e0216890. <https://doi.org/10.1371/journal.pone.0216890>

- Tatishvili, M.R., Mkurnalidze, I.P., Samkharadze, I.G., Chinchaladze, L.N., n.d. Application of Satellite Imagery in Forestry for Georgia 8.
- Wang, F., Gao, J., Zha, Y., 2018. Hyperspectral sensing of heavy metals in soil and vegetation: Feasibility and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing* 136, 73–84. <https://doi.org/10.1016/j.isprsjprs.2017.12.003>
- Whitlock, B.A., Baum, D.A., 1999. Phylogenetic Relationships of *Theobroma* and *Herrania* (Sterculiaceae) Based on Sequences of the Nuclear Gene Vicilin. *Systematic Botany* 24, 128. <https://doi.org/10.2307/2419544>
- Xiong, D., Chen, J., Yu, T., Gao, W., Ling, X., Li, Y., Peng, S., Huang, J., 2015. SPAD-based leaf nitrogen estimation is impacted by environmental factors and crop leaf characteristics. *Sci Rep* 5, 13389. <https://doi.org/10.1038/srep13389>
- Yu, K., Anderegg, J., Mikaberidze, A., Karisto, P., Mascher, F., McDonald, B.A., Walter, A., Hund, A., 2018. Hyperspectral Canopy Sensing of Wheat *Septoria Tritici* Blotch Disease. *Front. Plant Sci.* 9, 1195. <https://doi.org/10.3389/fpls.2018.01195>
- Zarcotejada, P., Berjon, A., Lopezlozano, R., Miller, J., Martin, P., Cachorro, V., Gonzalez, M., 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, 271–287. <https://doi.org/10.1016/j.rse.2005.09.002>
- Zhang, Z., Liu, M., Liu, X., Zhou, G., 2018. A New Vegetation Index Based on Multitemporal Sentinel-2 Images for Discriminating Heavy Metal Stress Levels in Rice. *Sensors* 18, 2172. <https://doi.org/10.3390/s18072172>
- Zhao, X., Wang, W., Ni, X., Chu, X., Li, Y.-F., Sun, C., 2018. Evaluation of Near-Infrared Hyperspectral Imaging for Detection of Peanut and Walnut Powders in Whole Wheat Flour. *Applied Sciences* 8, 1076. <https://doi.org/10.3390/app8071076>
- Zhou, W., Zhang, J., Zou, M., Liu, X., Du, X., Wang, Q., Liu, Yangyang, Liu, Ying, Li, J., 2019. Prediction of cadmium concentration in brown rice before harvest by hyperspectral remote sensing. *Environ Sci Pollut Res* 26, 1848–1856. <https://doi.org/10.1007/s11356-018-3745-9>
- Zug, K.L.M., Huamaní Yupanqui, H.A., Meyberg, F., Cierjacks, J.S., Cierjacks, A., 2019. Cadmium Accumulation in Peruvian Cacao (*Theobroma cacao* L.) and Opportunities for Mitigation. *Water Air Soil Pollut* 230, 72. <https://doi.org/10.1007/s11270-019-4109-x>