Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

ASSESSING SPECTRAL SIGNATURES OF POTATO PLANTS INFECTED WITH POTATO VIRUS Y

by

Lloyd Michael Griffel

A thesis submitted in partial fulfillment of the requirements for a

Master of Science in Geographic Information Science

Idaho State University

Fall 2016

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Lloyd Michael Griffel find it satisfactory and recommend that it be accepted.

_____

Donna Delparte, PhD, Assistant Professor,
Department of Geosciences, Idaho State University
Major Advisor

_____

Carrie Bottenberg, PhD, Assistant Lecturer,
Department of Geosciences, Idaho State University
Committee Member

_____

John Edwards, PhD, Assistant Professor,
Department of Informatics and Computer Science,
Idaho State University
Graduate Faculty Representative

Acknowledgements

I would first like to thank my thesis advisor and mentor, Dr. Donna Delparte, associate professor in the Department of Geosciences at Idaho State University, for her incredible support, patience and guidance. She continually steered me in the right direction and provided exceptional coaching and expertise that I will always cherish.

I would also like to thank Dr. John Edwards, professor in the Department of Computer Science at Idaho State University, for his support in the computer modeling work and for providing the scripting to run the Support Vector Machine classifier.

I would also like to thank Dr. H. Carrie Bottenberg, program coordinator and assistant lecturer in the Department of Geosciences at Idaho State University for her guidance and support for this research.

I would also like to thank Denny Smith, field manager at Potandon Produce L.L.C. and former field inspector with the Idaho Crop Improvement Association, for his work and assistance in visually identifying PVY-infected plants and collecting field spectrometer data.

Finally, I want to thank my wife, Lori Griffel, and my children, Mikaela, Ethan, Jonas, and Rylen, for their unending support while attending graduate school and conducting this research. They always inspire me to be better in everything that I do and this work would not have been possible without their support.

Lloyd Michael Griffel

Table of Contents

List of Figures

List of Tables

ASSESSING SPECTRAL SIGNATURES OF POTATO PLANTS INFECTED WITH POTATO VIRUS Y

Thesis Abstract – Idaho State University (2016)

In the potato industry, PVY has resulted in significant economic harm to farmers and has, at times, disrupted seed supplies to commercial growers, especially in varieties with good marketing attributes but high disease susceptibility such as Russet Norkotah. Commercial potato growers rely entirely on seed producers and certification systems to get disease-free seed as they have no recourse to mitigate seed-borne PVY after the seed is planted. However, seed growers and certification agencies are currently unable to control PVY infection in the industry's seed pipeline and this has a significant impact on commercial markets and regional economies. In this study, it is shown that PVY-infected potato plants in an agricultural production field produce different spectral signatures than neighboring non-infected plants. Using machine learning or machine vision analysis such as support vector machine classifiers can differentiate spectral signature of PVY-infected and non-infected plants at an accuracy of 89.8 percent. This was achieved in a field showing significant crop canopy variability as identified by remote sensing and cluster analysis.

1 Introduction

1.1 Agricultural Sustainability

Modern agriculture worldwide is under increasing pressure to decrease inputs such as water, fertilizer, and pesticides, while balancing increased production demand as global population and urban centers grow. Agriculture will need to double yields by 2050 to meet global food demand (Tilman et al. 2011). Crop inputs such as soil, water, seed, and fertilizer are finite or even scarce depending upon the region, and this yield increase will likely have to occur without a corresponding increase in inputs. To produce higher yields with the same or less inputs, site specific agricultural applications (precision agriculture) will need to be improved. In some instances in the United States, farmers are already varying application rates of fertilizer, seed, and irrigation water within fields that were historically treated in a homogenous way by applying site-specific management practices (Schmitt 2002). This is made possible with site-specific or targeted soil sampling and specialized application equipment that can alter applications on the fly using GIS and GPS technologies. Spatial data management, geostatistics, and remote sensing are also being utilized to analyze and manage regional disease pressures with the goal of improving Integrated Pest Management (IPM) strategies that result in a decreased dependency upon pesticides (Nelson et al. 1999).

Remote sensing technologies and analytics utilizing very high-spatial resolution sensors attached to unmanned aerial systems (UAS) are playing a major role in the continuing evolution of precision agriculture, where plant or even leaf scale analytics will provide an agronomic and economic benefit to growers. One example of this would be the detection and mapping of Potato Virus Y (PVY)-infected plants in seed producing fields using spectral signatures. Of the many diseases impacting the potato industry, PVY is among the most severe and can be responsible for yield losses of up to 80% to 90% (Kerlan 2008; Nolte et al. 2004).

Among the many crops produced in the world, potatoes (*Solanum tuberosum*) are among the most important economically, especially in Idaho. Potatoes are indigenous to the mountainous Andes region of South America and were originally domesticated and cultivated by the Incas and later distributed throughout the worlds by Spanish explorers becoming an important food staple for Europe by the 1700s (Alvarez et al. 2003). Today, potatoes are produced in about 150 countries and are the world's fourth most important crop behind wheat, rice, and maize (Abdelhaq et al. 2006). In 2015, U.S. farmers planted over 1 million acres of potatoes valued at almost $4 billion dollars. In Idaho, farmers planted 325,000 acres generating almost $900 million in revenue ("National Agricultural Statistics Service" n.d.). One factor that makes potatoes different from many other crops is that the tuber (starch storage organ) is actually a modified underground stem with an outer skin, vascular tissue, storage tissue, and undeveloped leaf buds (eyes) on the surface. Potato tubers serve two primary purposes – they can either be consumed as a

food source by humans and some domesticated animals or used as seed to produce more potato plants (Alvarez et al. 2003).

Seed potato inspection and certification in Idaho is managed by the Idaho Crop Improvement Association (ICIA) (Alvarez et al. 2003). The nonprofit grower association was designated to administer seed certification by the Regents of the University of Idaho as per Senate Bill No. 107, the "Seed and Plant Certification Act of 1959" which was enacted by the Idaho Legislature. Seed potato producers are required to meet certification standards regulating multiple disease thresholds in order to sell certified seed that is of high quality and free of disease. The association conducts multiple visual field inspections of every seed potato production lot during the growing season and laboratory tissue testing of samples collected during field inspections in an effort to quantify diseases including plant viruses. Growers are also required to submit tuber samples from each lot after harvest for the winter grow out test in Hawaii where the tubers are planted and subsequent plants are inspected and tested for multiple viral pathogens and other issues impacting seed quality. If seed potato lots exceed specific tolerances, marketing and plant-back restrictions are put in place that can severely limit that lot's marketability triggering economic losses to producers and disrupting seed supplies for commercial producers. A significant challenge seed certification agencies face is a lack of a wide-scale and accurate detection system to identify and quantify PVY-infected potato plants in seed production fields. Current programs rely on visual inspections during the growing season and post-harvest tuber samples that may not accurately represent actual seed lot

infection levels. These inconsistencies can often lead to certification errors that may show seed lots with artificially low or high infection levels.

There are no known cures or methods to directly control PVY infection in potato plants and commercial potato growers must plant PVY-free seed to maximize yields and quality (Strand and Rude 2006). Seed potato growers typically rely heavily on insecticide applications to suppress insect vectors known to spread viruses like PVY from plant to plant during the growing season (Abdelhaq et al. 2006). However, pesticide chemistries are becoming more limited because of unintended negative environmental consequences and the harmful effects some insecticides can have on beneficial organisms such as bees or aphid predators. More modern formulations are selective in toxicity and target specific insects like aphids while not harming beneficial organisms. However, these chemistries often do not immediately kill aphid targets which provides a window for continued disease transmission since an aphid carrying PVY can inoculate a plant in less than one minute (Whitworth, Nolte, et al. 2010). Mineral oil applications to potato plant foliage have also been used with limited success by creating a physical barrier that reduces aphid stylet penetration but the efficacy of the application depends upon coverage and can also result in crop damage from the phytotoxic effects oils have on leaf cell structure (Dupuis et al. 2014).

Seed growers also utilize cultural practices by mechanically removing infected plants with “rogueing crews” consisting of workers walking through fields visually inspecting plants for known virus symptoms and mechanically removing suspect plants.

The effectiveness of rogueing is highly variable as individual crew member training, experience, ambient light, crop and field conditions, and varieties can impact whether or not visual symptoms manifest in a way discernable to the human eye. For example, Russet Norkotah and Shepody varieties exhibit very mild or latent symptoms of PVY infection but will still suffer significant yield losses (Nolte et al. 2004; Whitworth, Nolte, et al. 2010). Some growers have also over-applied key crop nutrients such as nitrogen in the hopes that excessive vine growth would overcome PVY-induced yield losses. However, research has shown that adding extra fertility such as nitrogen does not mitigate yield drag caused by PVY (Whitworth et al. 2006). This has very detrimental effects on grower profit margins as well as contributing to groundwater contamination and other negative environmental impacts from the over-application of crop nutrients. Unfortunately, these regulatory-, pesticide-, and agronomy-based measures have not eradicated viruses such as PVY from potato seed producing operations. If anything, PVY is becoming more endemic and is now considered by many as one of the most significant threats to seed potato producers and their commercial customers on a global scale (Nolte et al. 2002).

1.2 Broader Impacts

In Idaho, PVY infection in potatoes plants has resulted in significant economic harm to farmers and has, at times, disrupted seed supplies to commercial growers,

especially in varieties with good marketing attributes but high disease susceptibility such as Russet Norkotah (Whitworth, Nolte, et al. 2010). PVY also produces mild or latent symptoms in Russet Norkotah making it more difficult to visually detect (Whitworth, Hamm, et al. 2010). Commercial potato growers rely entirely on seed producers and certification systems to get disease free seed as they have no recourse to mitigate seed-borne PVY after the seed is planted (Alvarez et al. 2003). However, seed growers and certification agencies currently struggle to control PVY infection in the industry's seed pipeline and this has a significant impact on commercial markets. Some of the most recent economic data indicates PVY reduces total potato production in Idaho by about 2.3 million hundred weight (cwt) annually resulting in an approximate loss of $14 million, depending upon current market pricing. The direct cost of PVY to the Idaho economy is about $19.5 million and economic modelling indicates the total impact exceeds $33 million annually (McIntosh et al. 2015).

A new way of detecting and quantifying PVY-infected plants that is more accurate than the current industry practices and standards could provide farmers and certification agencies a means to significantly reduce PVY levels within the industry's seed pipeline and restore lost yields and profits to farmers and the economy as a whole. A technological approach that will give seed growers the ability to destroy infected plants thereby reducing PVY inoculum and provide better seed potato quality to commercial customers will contribute to increased profits and improved environmental sustainability. By lessening dependence upon pesticides and maximizing crop inputs such as fertility,

irrigation, and land resources, this technology and analytics will help seed potato growers produce a healthier and more vigorous strains of seed stock. In turn this will help commercial growers achieve higher yields and finished product for every unit of input applied to fields.

1.3 Research Goals and Objectives

Spectral data collected with UAS mounted sensors and a field spectrometer (Analytical Spectral Devices, a PANalytical Company) from infected and healthy plants in a potato production field in eastern Idaho during the 2015 growing season are analyzed in this study to determine the type and structure of spectral differences with the goal of developing a methodology to accurately identify and map individual plants with PVY using remote sensing methodologies. A field with a typical crop rotation for eastern Idaho and managed with typical production practices is also analyzed to determine its heterogeneous features. PVY-infected plants in the field during the 2015 growing season were visually identified and assessed for spectral reflectance properties and the data subjected to machine-learning algorithms to look for potential spectral signatures. The research goal in chapter 3 is to evaluate the heterogeneity of field characteristics that leads to crop variability. The hypothesis being tested is that a selected field in southeastern Idaho is heterogeneous relative crop growth. Chapter 4 will focus on the hypothesis that there is a significant difference in how electromagnetic energy from the

blue to NIR interacts and reflects from a PVY infected plant versus a non-infected plant. Both hypotheses are important research goals since crop heterogeneity may contribute to variations in spectral responses of PVY-infected and healthy plants that influence the accuracies of classification efforts.

2 Literature Review

2.1 Modern Potato Production

From an agronomic perspective, potatoes are different from many other crops in that they are not commercially produced from seeds (Abdelhaq et al. 2006). Although potato plants can produce flowers during the growing season, they often drop before pollination meaning true potato seeds are very rare and also genetically unstable. Some varieties such as the Russet Burbank, one of the most common planted varieties, is male sterile as it does not produce viable pollen (Alvarez et al. 2003). Outside of breeding programs where true potato seeds are used to propagate new genetic hybrids, true potato seeds are not a viable seed source for most potato growing operations. Instead, potato tubers are typically used to promulgate the plants into the next year as a form of nonsexual vegetative propagation by producing new stems from the eyes when conditions are favorable for growth (Alvarez et al. 2003). During the growing season, potato plants progress through five growing stages: sprout development, vegetative growth, tuber initiation, tuber bulking, and maturation (Figure 1).

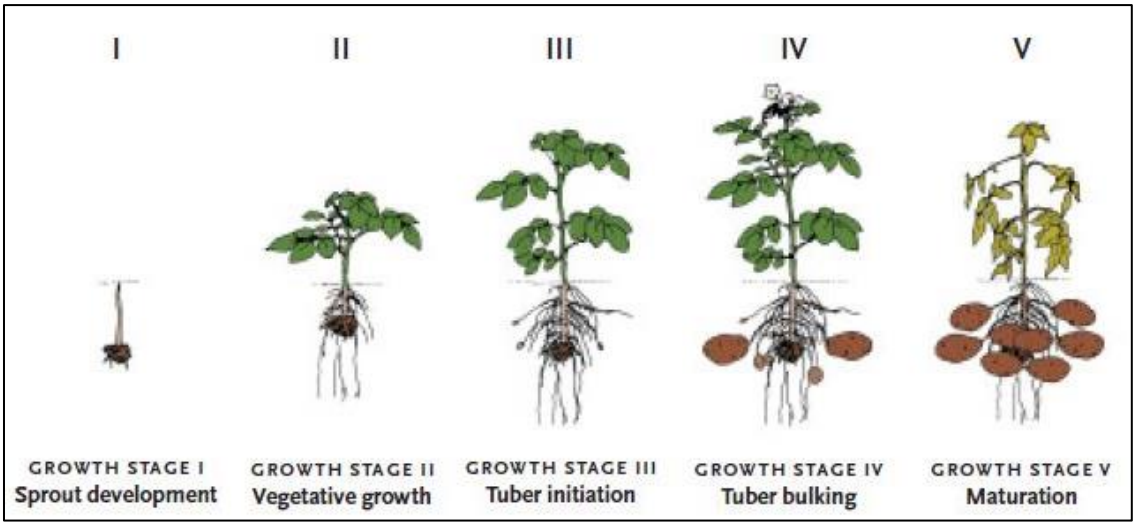

Figure 1. The five growth stages of a potato plant during the growing season are illustrated above. Potatoes are unique when compared to most other crops in that they are not produced from true seeds but are seeded with tubers produced from the previous season (adapted from Alvarez et al. 2003*).*

The fact that potatoes are produced via vegetative propagation differentiates potato producers into two typical categories: seed or commercial operations. Seed potato growers produce potato crops with the primary purpose of marketing their crop to commercial growers as "mother" seed for the commercial crop. Commercial operations then produce a potato crop where the "daughter" tubers are harvested to be sold in table stock, dehydrated, and frozen markets.

The advantages of vegetative propagation include improved crop management and agronomy since there is very little or no genetic diversity from plant to plant and plant growth is vigorous producing high yields of uniform produce(Abdelhaq et al. 2006). However, a major disadvantage of vegetative propagation is that potato seed tubers are vectors for many bacterial, fungal, and viral pathogens which can be carried into

subsequent generations via seed tuber tissue (Strand and Rude 2006). Therefore many countries and states have set up strict seed potato certification standards and programs with the goal of limiting diseases such as viruses that can be carried over from year to year in seed stocks (Abdelhaq et al. 2006). Seed potato production usually starts with disease free stock from plantlets derived from clean and thoroughly tested tissue culture stock and propagated for 3 to 5 years in isolated areas before being used to seed commercial production fields (Halterman et al. 2012). Typically, one pound of seed tubers will produce 10 to 15 pounds of daughter tubers each year. Some growing areas in Idaho as well as the entire state of Montana are restricted to only seed potato production and depend upon geographic segregation from commercial potato crops which typically carry more disease because of accumulated years of potential disease infection that could pass to nearby seed potato producing fields via insect vectors. However, since seed potatoes require multiple years of propagation within the seed pipeline, bacterial, fungal, and viral pathogens can infect and grow within the seed crop. Without intensive disease management programs, seed stocks can reach 100% disease infection levels within a few years (Halterman et al. 2012). It is not economically viable for commercial producers to use early generation seed stock as it is limited in quantity and much more expensive than later generation seed stock that has been propagated for 4 or 5 years.

2.2 PVY

PVY is a virus and member of the potyvirus genus that is readily transmitted (non-persistently) by multiple species of aphids and consists of multiple strains including PVY^{O}, PVY^{N}, PVY^{C}, and PVY^{NTN} (Abdelhaq et al. 2006). PVY is a single-strand RNA virus that infects a wide range of plant species causing wrinkles, chlorosis, and chlorotic or necrotic spots on leaves (Kogovšek et al. 2011). All PVY strains can cause internal and/or external tuber necrosis but the severity depends upon the potato variety and virus strain (Baarlen et al. 2005). PVY is considered a high risk for worldwide potato production due to its genetic variability, ease of transmission, and effect on yield and quality (Kerlan 2008; Whitworth, Nolte, et al. 2010).

Although countries and states have set up regulatory certification programs that include field inspections and extensive plant and tuber tissue testing, PVY levels continue to persist and in some cases increase depending upon disease pressure and the fact that many varieties, such as Russet Norkotah and Shepody mask PVY symptoms making it difficult to detect with current management strategies (Dupuis et al. 2014; Nolte et al. 2004; Whitworth et al. 2006, 2010). Many sources have documented significant yield losses from PVY as well as direct economic harm to seed and commercial potato growers as well as in-direct economic harm to agricultural industry partners that depend upon potato production. For instance, in three common potato varieties, Russet Burbank, Russet Norkotah, and Shepody, PVY has been shown to reduce farmer gross revenue by about $17/ha for every one percent of infection (Nolte et al. 2004). Improved line

selections of the Russet Norkotah variety, such as RN3, that result in increased vine vigor, are more prone to PVY infection (Whitworth, Hamm, et al. 2010). Recent economic data indicates each percent of PVY-infected plants in a fresh market Russet Burbank crop can reduce grower returns by $5.13 to $18.06 per acre and by $4.26 to $14.08 in processing crops (McIntosh et al. 2015).

2.3 Remote Sensing of Vegetation and PVY

Plant leaf structure has significant influence on leaf-light interactions (Figure 22). Although not identical in all plants, leaves consist of multiple layers of differing cells structures starting with the upper epidermis and lower epidermis with the lower consisting of stomata to facilitate air movement into the interior of the leaf. Below the upper epidermis is the palisade tissue containing chloroplasts which house chlorophyll and other photosynthesis pigments. Below that, mesophyll cells facilitate oxygen and carbon dioxide exchange. Chlorophyll facilitates the absorption of red and blue light to power photosynthesis while reflecting green wavelengths. Infrared wavelengths, which pass through the cuticle and epidermis, are reflected by mesophyll cells (Campbell 2007).

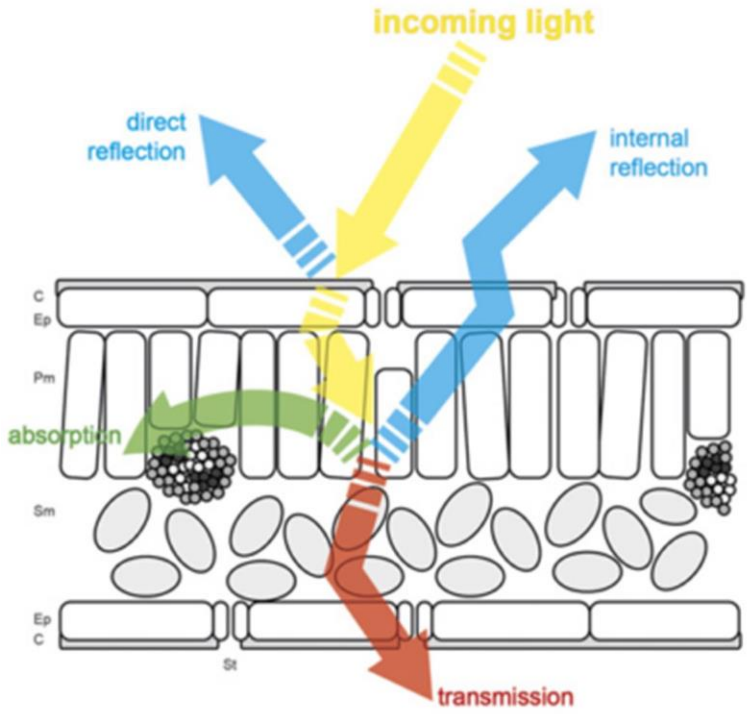

Figure 2. The diagram above shows a typical leaf cell structure and how light interacts with various leaf structures. Various pigments absorb blue and red wavelengths to provide energy for photosynthesis while reflecting green wavelengths. Internal mesophyll cells either reflect or transmit infrared wavelengths, very little of which is absorbed by plant cells (adapted from A.-K. Mahlein 2016*).*

It has been documented that PVY virus particles and inclusion bodies, the cylindrical protein bodies produced via expression of the viral genome, have been found in epidermis and mesophyll cells and in vascular tissues of infected plants (Otulak and Garbaczewska 2012). Previous studies indicate the virus affects expression of multiple genes including those involved with photosynthesis as well as several stress-response genes that produce heat shock and wound inducing proteins (Pompe-Novak et al. 2006).

2.4 Precision Agriculture

Agricultural production fields exhibit in-field heterogeneity that can cause significant yield and crop quality variation (Patzold et al. 2008). It is understood that soil heterogeneity in relation to parent material, climate, topography, and microbial and vegetative populations greatly influence fertility, crop water usage, and disease and pest pressure and the overall plant-by-plant response to biotic and abiotic stresses (Adamchuk et al. 2010). Although difficult to truly understand the origins of site-specific farming or precision agriculture, the modern nomenclature, farmers have been utilizing the concept for generations to manage crop inputs more accurately. For example, subsistence farming was carried out on very small areas of land that were naturally segregated for seeding, fertility, and sometimes irrigation by known or visual differences in plant growth. Today, large farming operations with fields covering large areas use grid or zone methodologies and modern equipment equipped with global positioning systems (GPS) to vary seed, fertility, and sometimes irrigation inputs to better match the varying needs within the field (Oliver 2010; Patzold et al. 2008; Schmitt 2002). The advent of remote sensing and geographic information systems (GIS) has added additional levels of data and analytics to make farming more efficient and sustainable. A key biophysical metric, leaf area index (LAI), that has been used model yield predictions, showed strong correlations to multiple vegetation indices (VIs) for corn and potato crops including the normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), transformed soil-

adjusted vegetation index (TSAVI), modified soil adjusted vegetation index (MSAVI2), and perpendicular vegetation index (PVI) with the MSAVI2 having the strongest correlation and dynamic range (Wu et al. 2007). It has also been shown that VIs exhibited strong correlations to yield with MSAVI2 having the most consistent correlation values (Yang and Everitt 2012). Remote sensing analysis has also been used to accurately detect and map sugarbeet plants with symptoms of beet cyst nematode and Rhizoctonia crown and root rot (Hillnhütter et al. 2011).

2.5 Advanced Precision Agriculture and Unmanned Aerial Systems

With the recent advancements in Unmanned Aerial Systems (UAS) and robust compact sensors, farmers are starting to have access to more advanced data and analytics. UAS offers multiple advantages over aircraft- or satellite-borne sensors including flight schedule and sensor flexibility, cost, and spatial resolution (Zhang and Kovacs 2012). This provides a new platform for remote sensing and to evaluate and study disease incidence in individual plants or even individual leaves (A. K. Mahlein et al. 2010). An example of this is found in research showing that changes in reflectance in certain regions of the electromagnetic spectrum relative to leaf vascular tissue in potatoes infected with Potato Yellow Vein Virus (PVYY) provided early detection of infection before symptoms of chlorosis where apparent to the trained eye (P. Chávez et al. 2009). Research also showed very high accuracies when classifying various plant diseases using

Support Vector Machine (SVM) classifier including those induced by fungal, viral and bacterial pathogens (D.Pujari et al. 2016). Originally introduced as the Support-Vector Network, SVM is a machine-learning binary classifier where input vectors are non-linearly mapped to a high dimension feature space where a linear decision surface is constructed (C.~Cortes et al. 1995).

3 An Analysis of Field Heterogeneity

3.1 Introduction

Understanding, measuring, and mapping field heterogeneity continues to be a major challenge for farmers moving to precision agricultural practices. Implementing variable rate technologies (VRT) to improve efficiencies of fertilizer, seed, and/or irrigation water could have significant positive impacts on environmental sustainability, crop yields and quality, and farm profits by matching crop inputs to heterogeneous field variability but establishing the cause-and-effect relationships of the biotic and abiotic factors that cause field heterogeneity impacting crop growth continues to be a major challenge (Schmitt 2002).

Although agricultural production fields are typically treated in a uniform manner when it comes to fertility, irrigation, and crop protection applications, it is known and accepted that soil physical attributes and chemistry can and often do vary within a single

field which impacts crop performance and yields, fertilizer, irrigation, and crop protection product performance (Patzold et al. 2008). Physical and chemical soil properties are the result of thousands of years of soil-forming features including parent material, topography, climate, soil and plant organisms, and time which result has resulted in soil variability at regional, field, and within field scales (Adamchuk et al. 2010).

An important question when dealing with in-field heterogeneity is whether or not the spatial patterns of crop health and vigor are random in nature or are being influenced by other factors. Often, crop canopy data measured by remote sensing can be analyzed to determine the potential random or biased nature of crop canopies using advanced geostatistics (Peeters et al. 2015; Rey-Caramés et al. 2015). Remote sensing of agricultural crops has shown continued promise as an effective tool to measure crop performance and variability both spatially and temporally in ways that are more efficient and cost effective than on-the-ground field measurements. VIs derived from multi- or hyperspectral data have been shown to be very effective to show and measure within-field variation (Hatfield and Prueger 2010). LAI is a dimensionless measurement used to quantify a plant's green leaf area or canopy which correlates to its ability to capture solar radiation for photosynthesis and plant growth. Research shows that LAI in potatoes is very closely related to tuber yield where LAI and ground cover were highly correlated despite variances in management practices and LAI accounted for 74 to 79 percent of tuber yield (Boyd et al. 2002). Much research has been done on how to measure LAI via radiative transfer models or by using VIs. Radiative transfer models (RTMs) rely on

extensive input data that can be difficult to acquire on a broad field scale and are difficult to implement in production agricultural practices while VI-LAI relationships in potatoes have shown strong correlations, especially when using MSAVI2 (Wu et al. 2007). Research also shows that VIs derived from remote sensing data have a high correlation to field level potato yields (Bala and Islam 2009).The purpose of this study is to conduct an analysis of heterogeneity of a selected agricultural production field in southeastern Idaho to test the hypothesis that the field is heterogeneous relative to crop growth as measured by MSAVI2 processed from high spatial resolution satellite imagery data and that underlying spatial patterns are not random. Crop growth and yield variability will be extrapolated with remote sensing data from high spatial resolution multispectral satellite imagery. This study is important because the field is typical of many agricultural production fields in the Pacific Northwest and is also the site of an additional study focusing on developing spectral signature/s of potato plants infected with Potato Virus Y (PVY) and it is not known at this time as to whether or not crop heterogeneity could contribute to variations in plant spectral responses to the viral infection. This chapter answers the following research questions:

- Is the crop variability completely random?
- How do elevation and topographical features relate to crop variability?
- How does soil variability measured by electrical conductivity relate to crop variability?

3.2 Materials and methods

3.2.1 Field Site

The field site encompasses 150 acres and is located in southeastern Idaho at an altitude of about 1,285 meters (Figure 3). The field was chosen because of its crop rotation, standard irrigation system, and typical agronomic practices. Coordinates are not being provided because the grower has requested the field location remain anonymous.

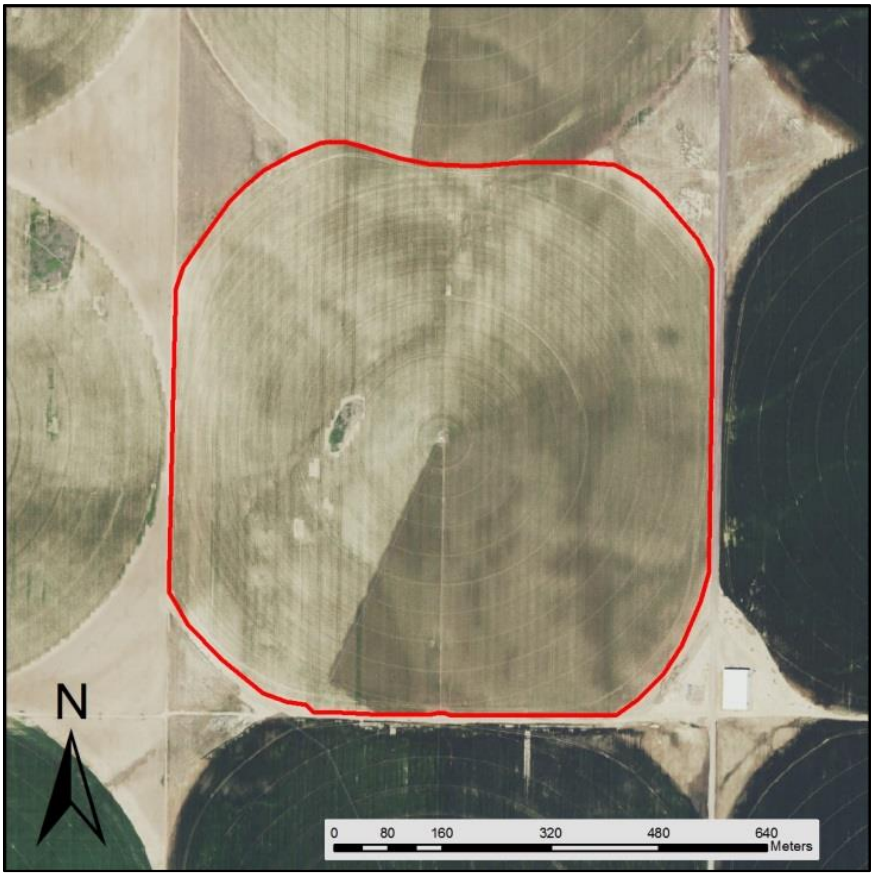

Figure 3. The red border bounds the farmed acres of a typical agricultural production field in southeastern Idaho as shown by 2015 NAIP (National Agriculture Imaging Program) imagery. Irrigation water is provided by a center pivot system to support a potato and cereal grain rotation. Coordinates are not being provided because the grower has requested the field location remain anonymous.

Southeastern Idaho predominantly consists of a high desert sagebrush steppe environment. The Snake River cuts through Snake River Valley that is used predominantly for agricultural production of potatoes, cereal grains, alfalfa, and sugarbeets. The area experiences a semiarid climate with below freezing winter temperatures, measurable spring and winter precipitation, and warm and dry summers. According to the Soil Survey Geographic (SSURGO) database (United States

Department of Agriculture, Natural Resources Conservation Service), the field soil consists of Pancheri silt loam, part of the coarse-silty mixed, frigid Xeric Haplocalcids classification split into subcategories based on topographical features (Figure 4).

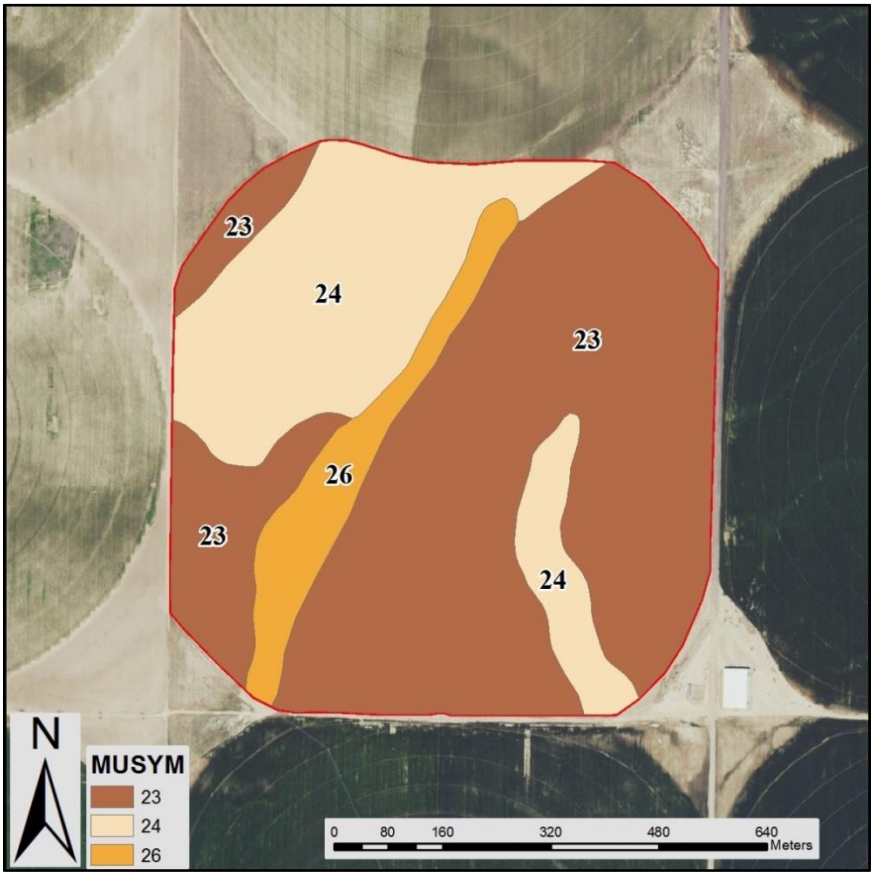

Figure 4: The Map Unit Symbol (MUSYM) key in the Soil Survey Geographic (SSURGO) database delineates the topographic breaks of the parent soil type, Pancheri, within the field site. The field soil consists of Pancheri silt loam, part of the coarse-silty mixed, frigid Xeric Haplocalcids classification.

Table 1 shows a breakdown of the field soil subcategories and associated acres. Table 2 shows a summary of the physical and chemical properties of the soil as provided by SSURGO.

Table 1. Soil subcategories for the field site derived from SSURGO.

Soil Name	MUSYM	Acres
Pancheri silt loam, 2 to 4 percent slopes	23	96.4
Pancheri silt loam, 4 to 8 percent slopes	24	43.0
Pancheri-Rock outcrop complex, 2 to 25 percent slopes	26	11.9

Table 2. Soil physical and chemical properties derived from SSURGO.

Depth (in)	Clay (%)	Available Water Capacity (in/in)	Organic Matter (%)	Cation Exchange Capacity (meq/100 g)	pH	Calcium Carbonate (%)	Gypsum (%)	Salinity (mmhos/cm)
0-6	8-17	0.19-0.21	1.0-2.0	7.0-18	7.4 - 8.4	1-5	0	0.0-2.0
6-10	5-18	0.19-0.21	0.5-1.0	4.0-16	7.9 - 9.0	10-30	0	0.0-2.0
10-17	5-18	0.11-0.13	0.0-0.5	4.0-15	7.4 - 9.0	15-25	0-5	2.0-8.0
17-33	5-18	0.11-0.13	0.0-0.5	4.0-15	7.4 - 9.0	15-25	0-5	2.0-8.0
33-60	5-18	0.11-0.13	0.0-0.5	4.0-15	7.4 - 9.0	15-25	0-5	2.0-8.0

The field site is managed with a typical potato and cereal grain rotation of one year planted to potatoes followed by two years planted to cereal crops and then back to potatoes the following year. It is irrigated with a center pivot system equipped with a corner extension arm with water supplied from an irrigation well supplied by the Eastern Snake River Plain Aquifer. For this study, the 2015 potato crop planted in the field site was analyzed to determine whether or not crop growth and canopy health are heterogeneous and not random spatially. Potato growth is classified into five growth stages: sprout development, vegetative growth, tuber initiation, tuber bulking, and maturation (Alvarez et al. 2003). The grower managed the field using a standard “flat-

rate" approach regarding fertility, planting, and irrigation and did not vary inputs to match in-field variability. This is still a common practice in production fields in Idaho.

3.2.2 Planting

The planting practices for the 2015 crop matched industry standards of other potato growers in the region. The field was seeded on April 29, 2015 and the variety planted was Russet Norkotah Line 278 with the premium quality tubers intended for the fresh table stock market and remaining off grade going to process channels. The seed consisted of generation 3 (G3) seed potatoes mechanically cut into pieces averaging 2.25 ounces/seed piece. An 8-row planter (Spudnik Equipment Company LLC, Blackfoot, Idaho) was used to plant the seed into rows spaced at 36 inches from center to center at an approximate 12 inch seed spacing approximately 6 inches deep (Figure 5). The planter was pulled with a tractor equipped with global positioning system (GPS) guided steering to ensure that rows were straight and in optimum alignment. Liquid injection systems were also attached to the tractor/planter to inject liquid fertilizer and soil fungicide products into the seed furrow during planting.

Figure 5. Seed potatoes are loaded with a belted elevator piler into the hopper of the eight-row potato seed planter. The planter was pulled with a tractor equipped with a GPS auto-steering system and equipped with liquid injection systems for fertility and fungicide products.

3.2.3 Agronomics

The fertility program for the 2015 potato crop at the field site was similar to other potato fields in the region. It consisted of multiple applications starting in the fall of 2014 after the grain crop was harvested. Approximately 50 pounds of nitrogen, 100 pounds of phosphate, 150 pounds of water soluble potash, 200 pounds of elemental sulfur, 5 pounds of zinc, and 10 pounds of magnesium were applied to every acre in the form of monoammonium phosphate (MAP) with an analysis of 11-52-00, K-Mag (The Mosaic Company) with an analysis of 00-00-22 11% Mg 22% sulfate, and Zinc Sulfate 35.5%

(Winfield Solutions LLC) with an analysis of 17.5% sulfur and 35.5% zinc. Additional liquid fertilizer was applied at planting through the liquid injection system on the planting equipment. The grower has requested specific product names be omitted to protect proprietary agronomic program information. This included 20 gallons/acre of ammonium phosphate with an analysis of 10-34-00, a 5 percent manganese solution at 0.6 gallons per acre, a 5 percent copper solution at 0.4 gallons per acre, 5 pounds of soluble boron per acre, and humic acid at 2 gallons per acre. An additional 75 pounds per acre of nitrogen in the form of UAN 32 with an analysis of 32-00-00 (45 percent ammonium nitrate, 35 percent urea, and 20 percent water) was applied during the season through the fertigation system on the center pivot irrigation system on a weekly basis based on in-season tissue sampling results.

Because of the intense fungal, insect, weed, and disease pressure for potatoes in southeastern Idaho, multiple crop protection applications occur starting with the seed cutting phase. The grower has requested rates not be published to protect proprietary agronomic program information. Right after seed pieces were cut, each was mechanically treated with a dry seed treat dust containing mancozeb (MZ) fungicide. As the seed pieces were planted, a liquid application system on the planter sprayed the soil surrounding the seed piece with Vertisan fungicide (DuPont) and Advise Four insecticide (Winfield Solutions LLC). Prior to plant emergence on June 3, 2015 during the sprout development stage, soil herbicides were applied to the exposed soil surface with a ground application unit and pushed into the soil profile approximately 1 to 2 inches deep with

irrigation water. The herbicide products included Sencor (Bayer Crop Science), Outlook (BASF), Prowl H2O (BASF), and Linex (NovaSource). Three in-season fungicide applications were applied during the month of July with an aerial applicator during the tuber bulking stage. Endura (BASF) was applied July 3, followed by a second application of Endura (BASF) and Roper (Loveland) on July 17. The last fungicide application took place on July 29 consisting of Tanos (DuPont). At the end of the growing season on August 28, 2015, a desiccant application consisting of Reglone (Syngenta) and Rely (Bayer Crop Science) was applied with a ground applicator to kill the above ground plant structure to force the tubers to mature and ripen for harvest and storage.

Because of the warm and dry semiarid high-desert environment, the field site required a significant amount of additional irrigation water to maximize the crop output. The first irrigation took place June 4, 2015 with a 0.5 inch application to water in the herbicide application. On June 8, the more intensive irrigation program began with an application of 1.5 inches as the crop had emerged and entered the vegetative growth stage and the need for water increased because of increasing evapotranspiration rates. The crop continued to receive approximately 1.5 inches per week until early August when crop water consumption began to decrease as the crop entered into the maturation growth stage (Figure 6). Irrigation application decisions were made on a weekly basis based on field scouting and soil moisture patterns measured by the irrigation manager. After vine kill, irrigation to the field site was discontinued for approximately two weeks to facilitate above-ground plant desiccation. Additional irrigation water was applied prior to harvest,

which started on September 21, 2015, to facilitate optimum soil moisture for tuber digging conditions. Harvest operations were completed on September 24, 2015 for the field site.

Figure 6. The field site on August 1, 2015. The plants are starting to transition from tuber bulking to the maturation stage which typically results in senescence of the crop canopy.

3.2.4 Remote Sensing

Because of the large scale of the field site and the limited "on-the-ground" resources to map the field variability, high spatial resolution multispectral satellite imagery was used to measure the 2015 potato crop photosynthesis patterns both spatially and temporally. The remote sensing data was collected with the Pleiades 1A and 1B satellites (Airbus Defence and Space) providing blue, green, red and NIR bands at a 2.0 meter native spatial resolution and provide by the J.R. Simplot Company (Boise, Idaho). The Pleiades satellite system was chosen because of its high temporal, spatial, and radiometric resolution. Pixel depth at acquisition is 12 bits increasing the likelihood that subtle differences in canopy difference will be detected. The data was delivered as a standard Ortho product including World Geodetic System (WGS84) georeferencing, orthorectification using the Reference3D dataset (Astrium's Elevation30 Suite), and viewing angle correction. Although several image dates were collected, only three were appropriate for the field heterogeneity analysis as intermittent cloud cover and local weather patterns prohibited the use of the other imagery dates. The dates used for this study are June 08, 2015 (early vegetative growth), July 12, 2015 (tuber bulking), and August 25, 2015 (maturation) and the collection parameters are shown in table 3.

Table 3. Satellite imagery specifications.

Acquisition Date	Off-Nadir (degrees)	Red (nm)	Green (nm)	Blue (nm)	NIR (nm)
June 8, 2015	12.97	620-700	510-590	450-530	775-915
July 12, 2015	12.08	620-700	510-590	450-530	775-915
August 25, 2015	15.84	620-700	510-590	450-530	775-915

Initial data preprocessing including image calibration, atmospheric correction, and registration were executed with ENVI 5.3 (Harris Corporation) software. The raw data was radiometrically corrected using the Radiometric Calibration tool to convert the raw pixel digital numbers (DNs) to radiance values (µW/(cm2 * sr * nm). The Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) module was used for atmospheric correction and to convert the radiance values to surface reflectance values to improve the accuracies of temporal and spatial analyses (Hadjimitsis et al. 2010). Because of the timing and location of the imagery acquisitions, the Mid-Latitude Summer atmospheric and Rural aerosol models where used. The dark pixel reflectance ratio method to retrieve the aerosol amount and estimate the scene average visibility (Aerosol Retrieval) was not applied where the Pleiades multispectral data did not include any data in the 2,100 nm range. Following atmospheric correction, the corrected imagery data was reprojected to the Universal Transverse Mercator North American Datum 1983 (UTM NAD83) zone 12 north and registered to the 2015 National Agriculture Imaging Program county base map to ensure spatial alignment was as accurate as possible (Figure 7).

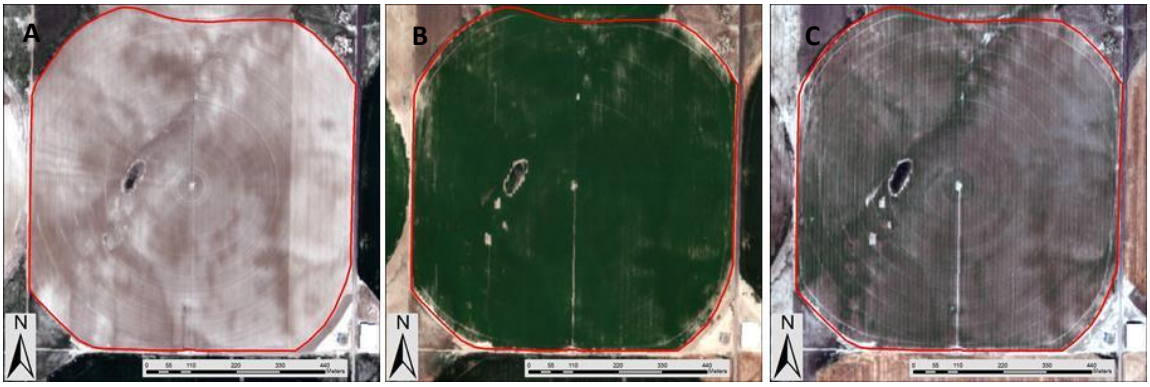

Figure 7. The three preprocessed satellite image tiles show the field site in natural color on June 8, 2015 (A), July 12, 2015 (B), and August 25, 2015 (C). The crop emergence pattern was captured on June 08, followed by the closed canopy on July 12, 2015 (middle), and August 25, 2015 (right). The crop emergence pattern was captured on June 08, followed by the closed canopy on July 12 and then crop maturation on August 25.

Following preprocessing, the analysis phase was conducted using ArcGIS 10.4 for Desktop (Environmental Systems Research Institute). To minimize any error associated with non-cropped portions of the field (rocky outcroppings, non-cropped areas, and irrigation pivot corners), a more accurate field boundary was digitized over the satellite imagery that excluded all areas within the field site that were not planted to potatoes. Using the preprocessed multispectral data, a MSAVI2 (1) image was generated for the three image dates to characterize and quantify the spatial and temporal crop vegetative patterns. The MSAVI2 equation, which is based on the SAVI equation that incorporates a soil adjustment factor (L) into the denominator of the NDVI equation, replaces L with a dynamic function that increases the dynamic range of the vegetation signal while minimizing soil background influences and also incorporates the near infrared (NIR) and red (RED) bands resulting in overall improved vegetation sensitivity (Qi et al. 1994).

$$MSAVI2 = \frac{(2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)}}{2} \quad (1)$$

3.2.5 Testing for Complete Spatial Randomness

Continuing in ArcGIS 10.4 for Desktop, the MSAVI2 raster data for the three dates was resampled to a five-meter spatial resolution using a bilinear technique. It was believed that further analysis would be computationally difficult with the native two-meter resolution. The raster data was then converted to a point feature class in order to begin to test for complete spatial randomness (CSR). During this analysis, the various processes produced z-scores and p-values to indicate whether or not the null hypothesis that the data was completely spatially random could be accepted or rejected. P-values represent the probability that the observed spatial features or values associated with the features are random in nature while the z-scores are the standard deviation values associated with a normalized distribution (Figure 8). Very low or high z-scores are found in the tails of the normal distribution and indicate that is it unlikely that the observed pattern is random. For this study, rejection of the null hypothesis would indicate the MSAVI2 values and their individual locations, which represent the continuous vegetation patterns in the field site, were not random but influenced by other spatial processes or phenomena.

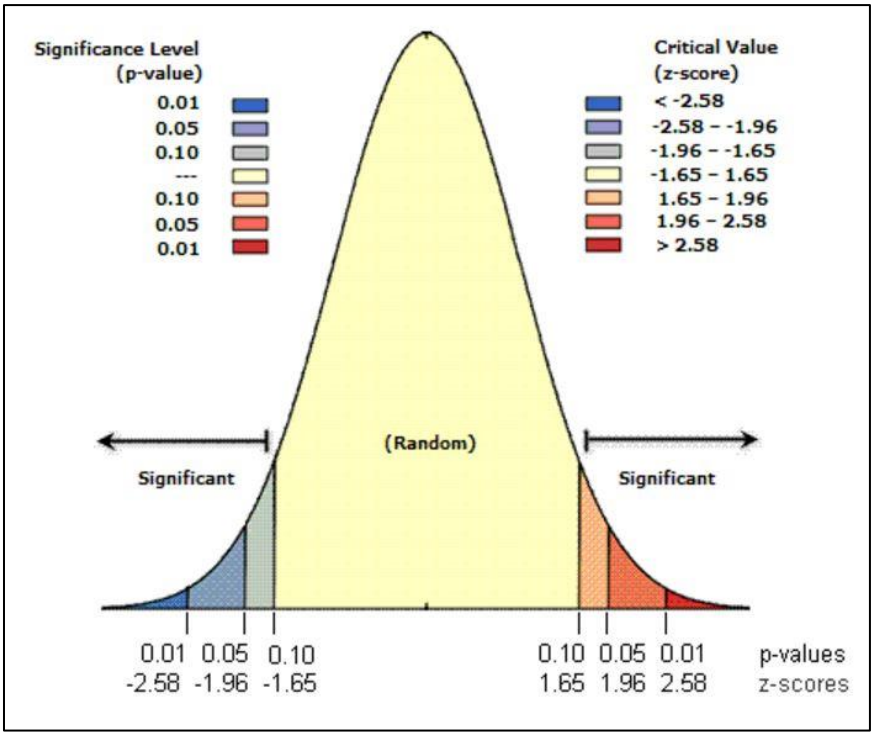

Figure 8. This graphical representation of the normal distribution pattern (ESRI) is used to assess whether or not features or values associated with features or random in nature. Significantly low p-values and significantly low or high z-scores indicate the null hypothesis that the data is random can be rejected.

As a starting basis, the point feature class data generated from the resampled five-meter spatial resolution MSAVI2 data was processed with the Average Nearest Neighbor tool in ArcGIS 10.4 for Desktop using Euclidian distance to measure the distance between each feature centroid and its nearest neighbor's centroid location and determine whether or not the point data spatial pattern was randomly distributed. The analysis parameter was confined to 592,562.78 square meters, the area of the crop canopy. Next, the feature class point data was interrogated with multiple iterations of the Incremental

Spatial Autocorrelation tool using Euclidian distance to determine the spatial distances where the z-scores or standard deviation peaked indicating the distance where data clustering was most pronounced. This process was computationally intensive given the large number of points and potential neighbors based on iterations of the number of distance bands chosen and the distance increment at which z-scores were calculated. Using the distances found with the Incremental Spatial Autocorrelation tool for each image date, the Spatial Autocorrelation tool was implemented using the Global Moran's I statistic to determine whether or not the MSAVI2 values and spatial patterns expressed clustering (Moran 2016). The inverse distance weighted (IDW) method was used to define the spatial relationship of the MSAVI2 point feature class given the continuous nature of the MSAVI2 values. Euclidian distance was used as the distance method and the distance where data clustering was most pronounced that was found during the Incremental Spatial Autocorrelation analysis was used as the distance band. The peak clustering determined from the June 8 and July 12 dates was used and compared across all image dates. If the results of the Global Moran's I analysis indicated a non-random pattern within the MSAVI2 data, a Cluster and Outlier Analysis using the Anselin Local Moran's I methodology to identify the significant high and low MSAVI2 values as well as any spatial outliers was performed (Anselin 1995). A Hot Spot Analysis using the Getis-Ord Gi* analysis was also conducted to identify statistically significant spatial clusters of high and low values (Getis and Ord 1992; Ord and Getis 2010). The inverse distance weighted (IDW) method was used to define the spatial relationship of the

MSAVI2 point feature class given the continuous nature of the MSAVI2 values. Euclidian distance was used as the distance method and the distance where data clustering was most pronounced for each date that was found during the Incremental Spatial Autocorrelation analysis was used as the distance band.

3.2.6 Exploratory Regression – Elevation Data and Derivatives

After testing the MSAVI2 point feature classes for the three growing season dates for CSR, an exploratory regression analysis was conducted to ascertain what topographical explanatory variables could explain the spatial variability of the crop canopy health measured by MSAVI2 for each image date. The analysis evaluated all possible combinations of explanatory variables using ordinary least squares (OLS) regression. It is believed topographical variations will likely result in differing soil types not measured or explained by the SSURGO data due to sampling limitations and vary the amount of incoming solar radiation the crop receives within the field which influences photosynthetic activity and plant biomass production. These topography based variations could result in in-field differences affecting crop canopy growth and health. The explanatory variables included elevation, aspect, slope, total incoming solar radiation (direct and diffuse), direct incoming solar radiation, diffuse incoming solar radiation, and the duration of direct incoming solar radiation for the growing season. The elevation data

and derivatives for the field site were derived from the United States Geological Survey (USGS) National Elevation Data 10 meter or better (NED10) provided by the USDA via the Geospatial Data Gateway (https://gdg.sc.egov.usda.gov/). The elevation data values were given in meters at a spatial resolution of approximately 8.9 meters.

Using ArcGIS 10.4 for Desktop to prepare the data for exploratory regression analysis, aspect and slope surface rasters were produced using the NED10 elevation data resulting in values for each pixel (Figure 9). Elevation was measured in meters, aspect in compass direction degrees, and slope in degrees.

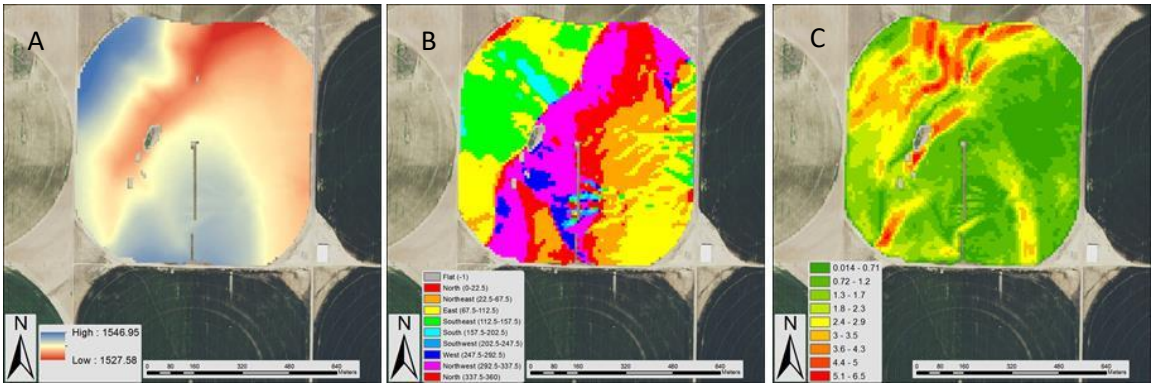

Figure 9. The tiles above display elevation (A), aspect (B), and slope (C) for the field site showing elevation in meters, aspect in compass direction degrees, and slope in degrees. The data was derived from USGS NED 10 meter data.

The incoming solar radiation surfaces were produced using the Area Solar Radiation tool with the time configuration specified for June 8, 2015 (crop emergence) through August 28, 2015 (crop canopy desiccation) (Figure 10). The solar radiation data is measured in watt hours per square meter, and the duration data is measured in hours.

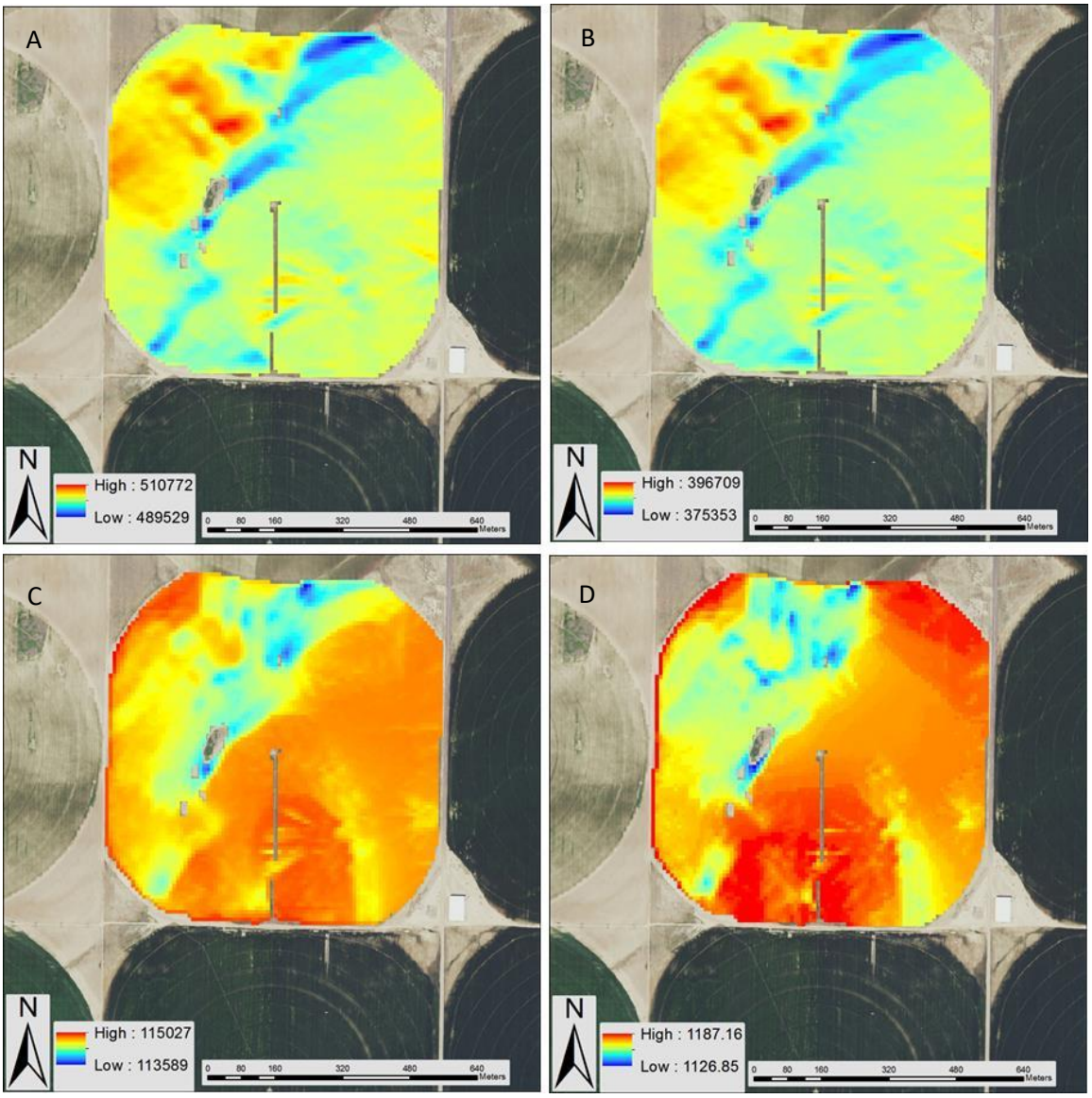

Figure 10. The tiles above display total incoming solar radiation (A), direct solar radiation (B), diffuse solar radiation (C), and duration of direct solar radiation (D) for the crop growing season from June 8, 2015 to August 28, 2015. The elevation data is measured in meters, the solar radiation data is measured in watt hours per square meter, and the duration data is measured in hours.

To conduct the exploratory regression analysis, the dependent and explanatory variables had to be combined into a single feature class where the corresponding variables shared a common spatial location. After completing the topography derivatives, a grid polygon feature class was constructed that matched the cell size and alignment of the elevation and topography raster data. The cell values of each topography derivative were extracted to the corresponding cell of the grid feature class attribute table. Using the Zonal Statistics tool, the mean values of the native resolution MSAVI2 rasters derived from the atmospherically corrected multispectral data where calculated for each cell of the grid polygon feature class. The resultant attribute table of the grid polygon feature class contained the corresponding elevation, topography derivatives, and mean MSAVI2 values for each cell at a spatial resolution of 8.9 meters (Figure 11).

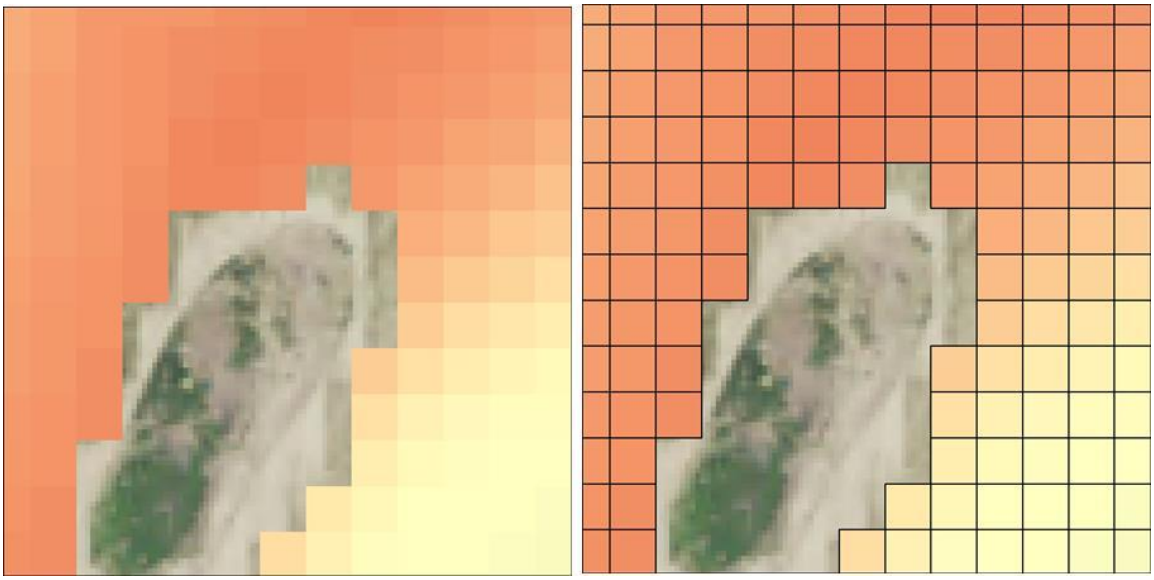

Figure 11. The left tile shows the elevation grid raster derived from USGS NED10 and the right tile shows the overlaying grid feature class matching the spatial resolution and alignment of the raster data. The attribute table of the grid polygon feature class con contained the corresponding elevation, topography derivatives, and mean MSAVI2 values for each cell at a spatial resolution of 8.9 meters.

Using the Exploratory Regression tool in ArcGIS 10.4 for Desktop, the elevation and topography derivatives (elevation, slope, aspect, total incoming solar radiation, diffuse solar radiation, direct solar radiation, and the duration of direct solar radiation) were used as explanatory variables against each MSAVI2 date (June 8, 2015, July 12, 2015, and August 25, 2015). The analysis looks for Ordinary Least Square models that best explain the dependent variable which, in this case, are the MSAVI2 variables. These explanatory variables where chosen because elevation data was readily available and the area is known for significant topographical changes within agricultural production fields. The analysis included modelling for single variables as well as all possible combinations

of the variables ranging from 2 possible combinations up to 7 possible combinations. For the purposes of this study, the following parameters are required to be considered a passing model: an adjusted R-squared value of 0.5, a maximum coefficient p value of 0.05, a maximum variance inflation factor of 7.5, a minimum acceptable Jarque-Bera p-value of 0.1, and a minimum acceptable spatial autocorrelation p value (based on the Global Moran's I test) for model residuals of 0.1.

3.2.7 Exploratory Regression – Soil Electrical Conductivity Data

An exploratory regression analysis was also conducted to determine whether or not soil apparent electrical conductivity (EC) readings could explain the crop canopy variability measured by the MSAVI2 analysis for the three image dates previously analyzed for complete spatial randomness. The EC data was provided for the study by the J.R. Simplot Company (Boise, Idaho) and was collected during the fall of 2014 with an EM38-MK2 proximal sensor measuring apparent conductivity in millisiemens per meter (mS/m) at a depth of 0.5 meters and 1 meter (Geonics Limited, Mississauga, Ontario, Canada) after the 2014 cereal grain crop was harvested. EC data has been shown to be effective when mapping the spatial variability of soil texture as influenced by soil moisture content, salinity, and clay which can significantly influence plant growth characteristics (Gooley et al. 2014; Guo et al. 2016; Mertens et al. 2008).

The EC data was mapped in transects spaced at approximately 16 meters and recorded in 1-second intervals with a Trimble Yuma 2 field computer linked to the EM38-MK2 sensor linked to a global positioning system (GPS) at 10 cm accuracy (Figure 12).

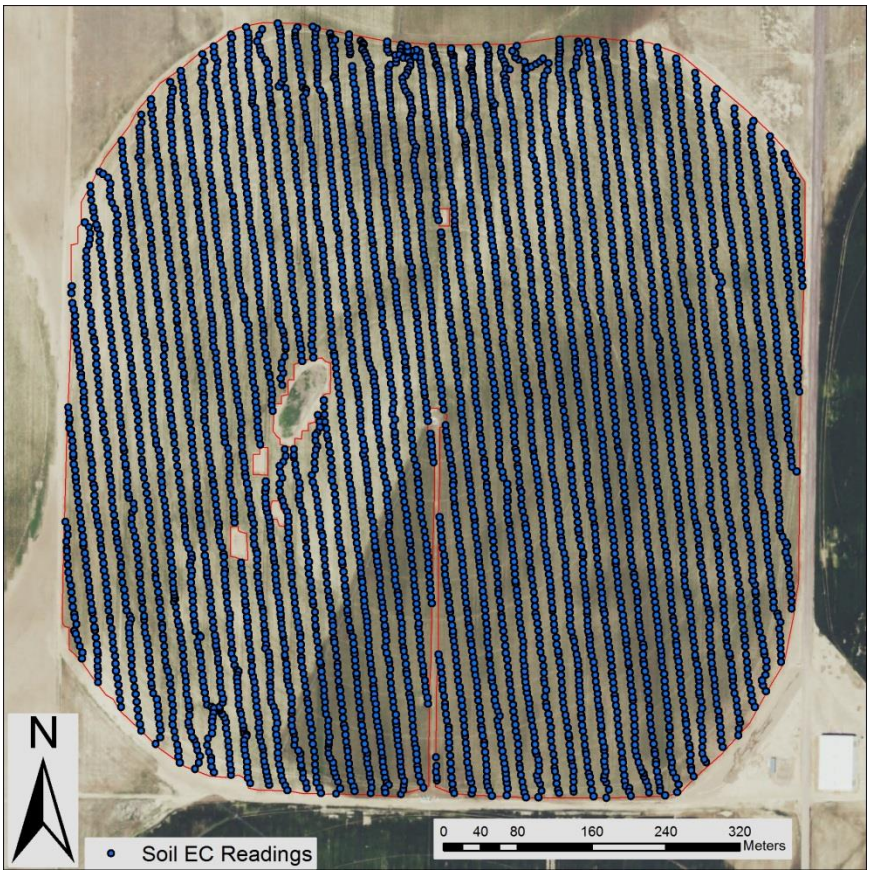

Figure 12. This map shows the 5,059 data points collected measuring apparent soil electrical conductivity readings (mS/m) in the field site. The data was provided by the J.R. Simplot Company and was collected during the fall of 2014.

The sensor was mounted in a proprietary sled designed to minimize temperature variations and dust exposure to the sensor and pulled with an all-terrain vehicle (ATV) with the operator using GPS-guided light-bar guidance for in-field navigation. A total of 5,059 data points were collected within the 2015 crop canopy field boundary used to process the remote sensing data earlier in this chapter (Figure 13). Table 4 shows the statistics for the 0.5 meter and 1 meter readings.

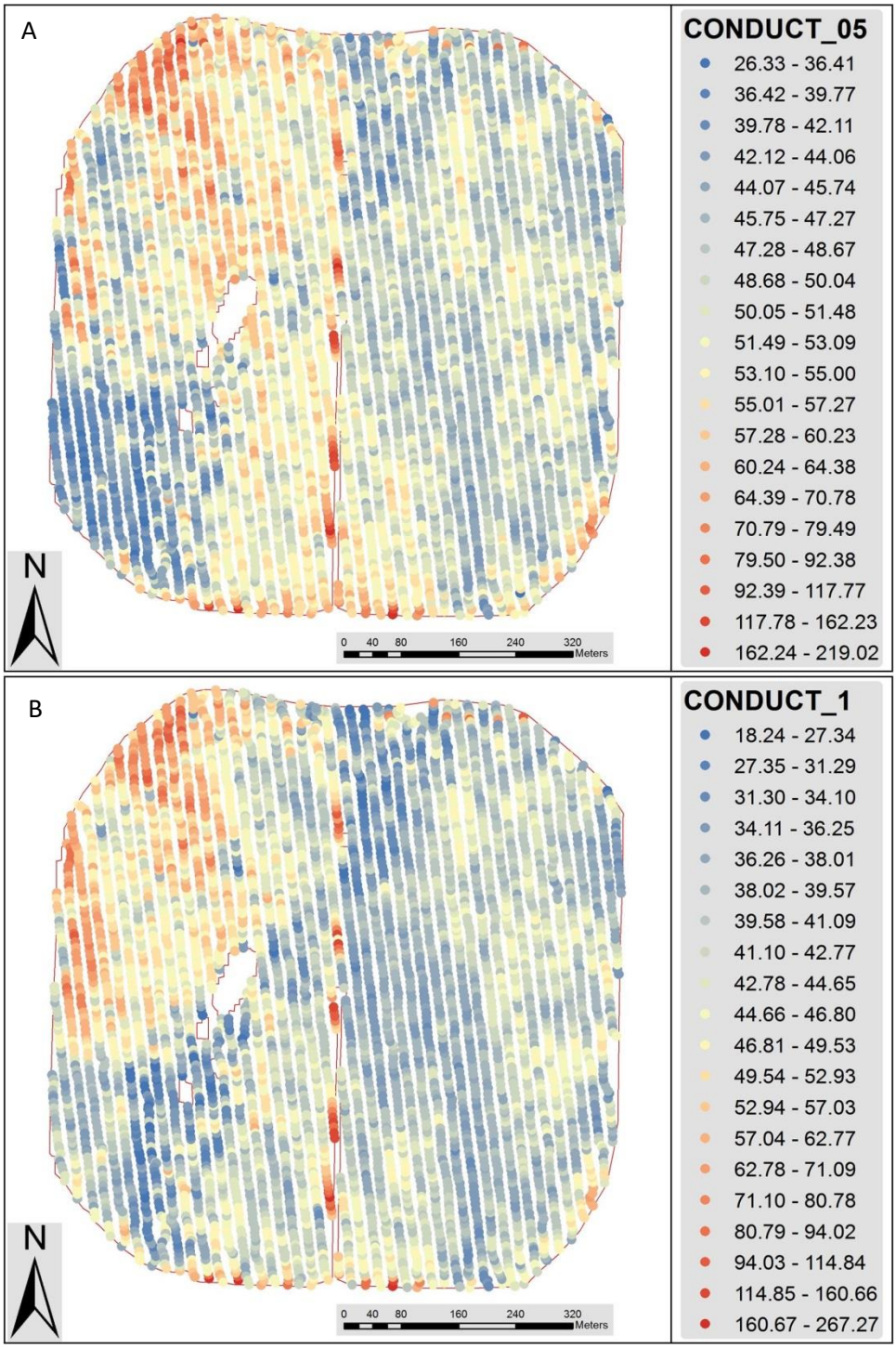
A
N
0 40 80 160 240 320
Meters
CONDUCT_05
26.33 - 36.41
36.42 - 39.77
39.78 - 42.11
42.12 - 44.06
44.07 - 45.74
45.75 - 47.27
47.28 - 48.67
48.68 - 50.04
50.05 - 51.48
51.49 - 53.09
53.10 - 55.00
55.01 - 57.27
57.28 - 60.23
60.24 - 64.38
64.39 - 70.78
70.79 - 79.49
79.50 - 92.38
92.39 - 117.77
117.78 - 162.23
162.24 - 219.02
B
N
0 40 80 160 240 320
Meters
CONDUCT_1
18.24 - 27.34
27.35 - 31.29
31.30 - 34.10
34.11 - 36.25
36.26 - 38.01
38.02 - 39.57
39.58 - 41.09
41.10 - 42.77
42.78 - 44.65
44.66 - 46.80
46.81 - 49.53
49.54 - 52.93
52.94 - 57.03
57.04 - 62.77
62.78 - 71.09
71.10 - 80.78
80.79 - 94.02
94.03 - 114.84
114.85 - 160.66
160.67 - 267.27

Figure 13. The maps above show the apparent soil electrical conductivity values in millisiemens per meter (mS/m) for the 0.5 meter (A) and 1 meter (B) soil depths of the field site as measured during the fall of 2014. The higher values could indicate soil with higher moisture, clay, and salt content as compared to areas with lower EC values.

Table 4. Soil electrical conductivity statistics at 0.5 and 1 meter depths.

depth	mean (mS/m)	minimum (mS/m)	maximum (mS/m)	std. dev (mS/m)
0.5 meters	51.308	26.328	219.023	9.069
1 meter	43.551	18.242	267.266	10.842

To prepare the EC data for exploratory regression, the Extract Multi Values to Points tool in ArcGIS 10.4 for Desktop was used to append the MSAVI2 values for the three image dates analyzed earlier for CSR at the exact location of every EC data point to the EC data point feature class. Using the Exploratory Regression tool, the EC data at 0.5 meters and 1 meter were used as explanatory variables against each MSAVI2 date (June 8, 2015, July 12, 2015, and August 25, 2015). For the purposes of this study, the following parameters are required to be considered a passing model: an adjusted R-squared value of 0.5, a maximum coefficient p value of 0.05, a maximum variance inflation factor of 7.5, a minimum acceptable Jarque-Bera p-value of 0.1, and a minimum acceptable spatial autocorrelation p value (based on the Global Moran's I test) for model residuals of 0.1.

3.3 Results

3.3.1 Remote sensing

Spatial maps showing the crop canopy variation as measured by MSAVI2 analysis were constructed for the June 8, 2015, July 12, 2015, and August 25, 2015 dates capturing early vegetative growth, tuber bulking, and crop maturation growth stages (Figure 14). The basic statistics for each image date are shown in Table 4.

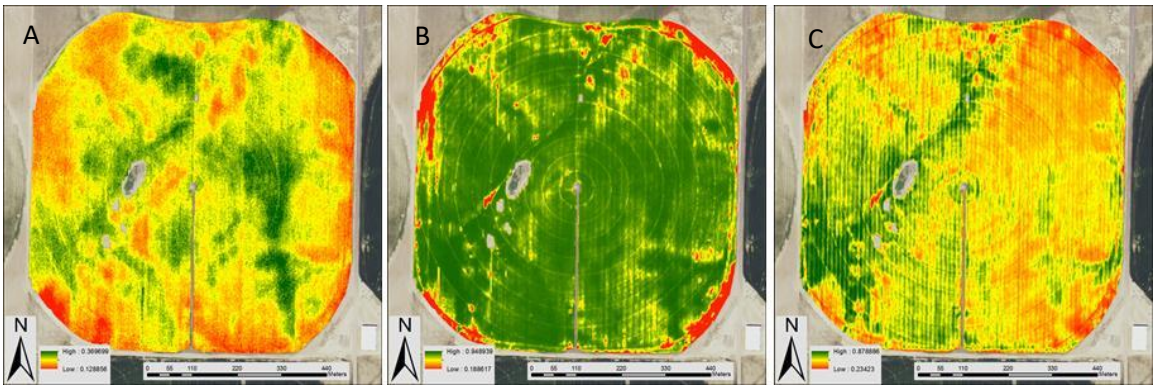

Figure 14. The MSAVI2 tiles for each date displayed as stretched rasters limited to 2.5 standard deviations on June 8, 2015 (A), July 12, 2015 (B), and August 25, 2015 (C). The crop emergence pattern was captured on June 08, followed by the closed canopy on July 12 and then crop senescence on August 25.

Table 5. MSAVI2 statistics for each image date.

Image Date	Mean	Min.	Max	Std. Dev.
June 8, 2015	0.258325	0.128856	0.369699	0.031612
July 12, 2015	0.898117	0.188617	0.948939	0.063283
August 25, 2015	0.545860	0.234230	0.878886	0.096557

3.3.2 Complete spatial randomness (CSR) analysis

The results of the average nearest neighbor spatial analysis for the MSAVI2 point feature classes derived from the five-meter resampled data for each date are shown in table 5. Based on the high z-score values and probabilities valued at zero, the null hypothesis that the spatial pattern of the MSAVI2 point feature classes for each date are random should be rejected.

Table 6. Average nearest neighbor results.

Image Date	Observed Mean Distance	Expected Mean Distance	Nearest Neighbor Index	z-score	p-value
June 8, 2015	5	2.5	2	294.65	0
July 12, 2015	5	2.5	2	294.65	0
August 25, 2015	5	2.5	2	294.01	0

The multiple iterations of the Incremental Spatial Autocorrelation tool on MSAVI2 point feature classes derived from the five-meter resampled data resulted in peaked z-scores for each image date (Figure 15). The June 8, 2015 data peaked at 106 meters based on 30 distance bands at five-meter increments with a beginning distance of six meters. The July 12, 2015 data peaked at 76 and 96 meters based on 30 distance bands at five-meter increments with a beginning distance of 6 meters. The August 25, 2015 data peaked at 351 meters based on 30 distance bands at 15-meter increments starting at 6 meters.

A B C

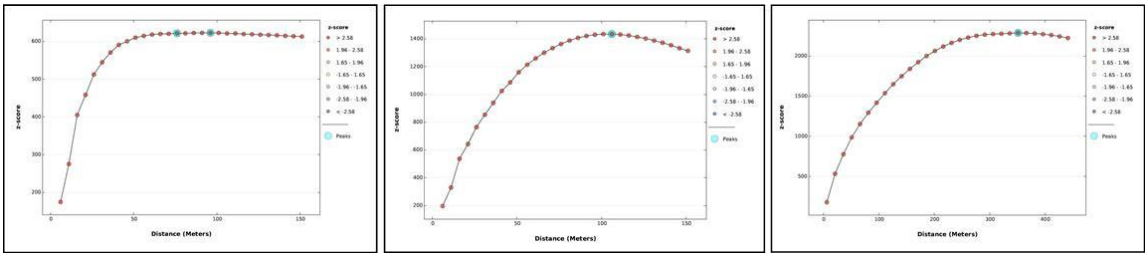

Figure 15. The three tiles above show the z-score peak location for the MSAVI2 point feature classes for June 8, 2015 (A), July 12, 2015 (B), and August 25, 2015(C). The results are based on multiple iterations of the number of increments and distance for each increment.

Using the spatial distances of 76 and 106 meters where the peak z-scores occurred for the June 8 and July 12 image dates, the Global Moran's I analysis was applied to each of the MSAVI2 point feature classes derived from the five-meter resampled data. The results are shown in table 6. Based on the high z-score values and probabilities valued at zero, the null hypothesis that the spatial distribution of the MSAVI2 values is randomly distributed should be rejected for each image date.

Table 7. Global Moran's I results.

Image Date	Moran's I Index	Expected Index	Variance	z-score	p-value
6/8/2015-76 meters	0.577167	-0.000042	0	1243.058092	0
6/8/2015-106 meters	0.483322	-0.000042	0	1379.225864	0
7/12/2015-76 meters	0.322644	-0.000042	0	695.292114	0
7/12/2015-106 meters	0.258456	-0.000042	0	737.983829	0
8/25/2015-76 meters	0.507308	-0.000042	0	1091.091986	0
8/25/2015-106 meters	0.454405	-0.000042	0	1294.816095	0

Where the Global Moran's I statistics indicated a nonrandom spatial pattern for the MSAVI2 values for each image date, the Anselin Local Moran's I analysis was

performed on each of the MSAVI2 point feature classes derived from the five-meter resampled data to identify spatial clusters of features with high or low values and spatial outliers at a 95 percent confidence interval (p-values less than 0.05). The output feature classes for each image date indicate the statistical standing of every feature and its corresponding MSAVI2 value showing whether each feature is part of a high or low cluster or a statistically high or low outlier compared to its neighbors (Figure 16).

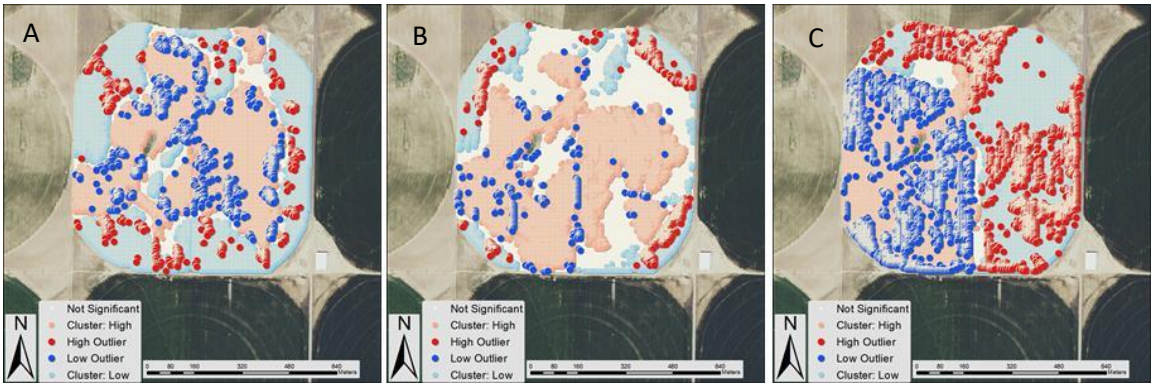

Figure 16. The tiles above show the resultant feature classes of the Anselin Local Moran's I analysis for the June 8, 2015 (A), July 12, 2015 (B), and August 25, 2015 (C) MSAVI2 image dates. Each image date had areas of both high and low clustering and outliers.

Table 8 shows the quantity of the areas of clustering and outliers for each image date in meters squared and total percentage of the field.

Table 8. Anselin Local Moran's I analysis results.

Image Date	Cluster: High (m^2)	Cluster: High (% of field)	Cluster: Low (m^2)	Cluster: Low (% of field)	High Outlier (m^2)	High Outlier (% of field)	Low Outlier (m^2)	Low Outlier (% of field)
June 8, 2015	195,550	33%	187,125	32%	11,350	2%	22,675	4%
July 12, 2015	218,700	37%	68,775	12%	15,425	3%	5,600	1%
August 25, 2015	173,275	29%	214,650	36%	50,700	9%	57,150	10%

A Hot Spot Analysis using the Getis-Ord Gi* analysis was also conducted to identify statistically significant spatial clusters of high and low values on each of the MSAVI2 point feature classes derived from the five-meter resampled data on the feature MSAVI2 values to identify the features that are spatially clustered by either high or low MSAVI2 values (Figure 17).

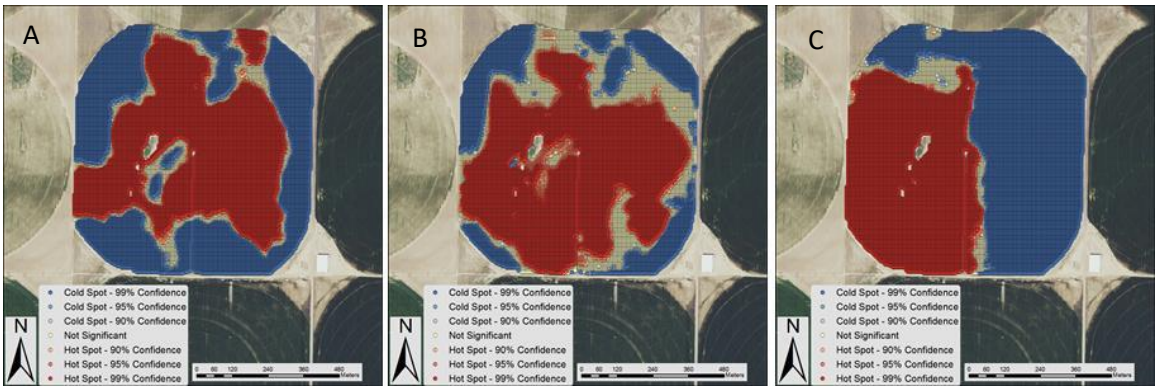

Figure 17. The tiles above show the resultant feature classes of the Getis-Ord Gi analysis for the June 8, 2015 (A), July 12, 2015 (B), and August 25, 2015 (C) MSAVI2 image dates. Each image date had areas of both high and low clustering at 90, 95 and 99 percent confidence intervals.*

Table 9 shows the quantity of the areas of clustering for each image date in meters squared and total percentage of the field.

Table 9. Getis-Ord Gi analysis results.*

Image Date	High: 99% CI (m^2)	High: 99% CI (% of field)	High: 95% CI (m^2)	High: 95% CI (% of field)	High: 90% of CI (m^2)	High: 90% of CI (% of field)
June 8, 2015	260,825	44.0%	13,600	2.3%	6,475	1.1%
July 12, 2015	300,975	50.8%	31,650	5.3%	15,625	2.6%
August 25, 2015	247,125	41.8%	5,650	1.0%	2,900	0.5%
Image Date	**Low: 99% CI (m^2)**	**Low: 99% CI (% of field)**	**Low: 95% CI (m^2)**	**Low: 95% CI (% of field)**	**Low: 90% of CI (m^2)**	**Low: 90% of CI (% of field)**
June 8, 2015	225,150	38.0%	12,275	2.1%	6,325	1.1%
July 12, 2015	104,925	17.7%	12,550	2.1%	7,900	1.3%
August 25, 2015	291,600	49.3%	4,875	0.8%	3,150	0.5%

3.3.3 Exploratory regression – Elevation Data and Derivatives

Table 10 shows the abbreviations for the applicable parameters to interpret all exploratory regression analysis using Ordinary Least Squares modelling. The results of the exploratory regression analysis are shown in tables 11 – 16 for the June 8, 2015 MSAVI2 values against the elevation and topography derivative explanatory variables (elevation, slope, aspect, total incoming solar radiation, diffuse solar radiation, direct solar radiation, and the duration of direct solar radiation). Only the top 3 regression results (highest adjusted R-squared result) for each possible number of combinations are included. Results were severe data multicollinearity or data redundancy occurred were not included.

Table 10. Table abbreviations for the exploratory analysis results.

AdjR2	Adjusted R-Squared
AICc	Akaike's Information Criterion
JB	Jarque-Bera p-value
K(BP)	Koenker (BP) Statistic p-value
VIF	Max Variance Inflation Factor
SA	Global Moran's I p-value
Model	Variable sign (+/-)
Model	Variable significance (* = 0.10, ** = 0.05, *** = 0.01)

Table 11. Highest adjusted R-squared results from the exploratory regression analysis of the elevation derivatives for the June 8, 2015 MSAVI2 values.

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.05	-31265.29	0.05	0	1	0	-ELEVATION***
0.03	-31118.27	0	0	1	0	-DIFFUSERADIATION***
0.03	-31046.78	0.01	0.07	1	0	-DIRECTDURATION***
0.08	-31478.73	0	0	1.03	0	-ELEVATION*** -ASPECT***
0.06	-31348.28	0.1	0	1.07	0	-ELEVATION*** -DIRECTDURATION***
0.06	-31345.69	0	0	2.32	0	-DIFFUSERADIATION*** -SLOPE***
0.1	-31640.25	0	0	1.3	0	-ELEVATION*** -ASPECT*** -SOLARRADIATION***
0.1	-31633.21	0.01	0	1.23	0	-ASPECT*** -DIRECTDURATION*** -SOLARRADIATION***
0.1	-31628.15	0.01	0	1.26	0	-ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION***
0.12	-31832.29	0	0	2.43	0	-ASPECT*** -DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***
0.12	-31818.02	0	0	2.5	0	-ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -SLOPE***
0.12	-31792.73	0	0	1.47	0	-ELEVATION*** -ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION***
0.13	-31900.48	0	0	3.12	0	-ELEVATION*** -ASPECT*** -DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***
0.13	-31898.62	0	0	3.21	0	-ELEVATION*** -ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -SLOPE***
0.13	-31874.28	0	0	862.94	0	-ASPECT*** +DIRECTRADIATION*** -DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***
0.13	-31900	0	0	10.59	0	-ELEVATION*** -ASPECT*** -DIRECTDURATION*** -DIFFUSERADIATION -SOLARRADIATION*** -SLOPE***
0.13	-31900	0	0	10.37	0	-ELEVATION*** -ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION** -SLOPE***

Table 12. Percentage of search criteria passed from the exploratory regression analysis of the elevation derivatives for the June 8, 2015 MSAVI2 values.

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	108	0	0
Max Coefficient p-value	< 0.05	108	99	91.67
Max VIF Value	< 7.50	108	89	82.41
Min Jarque-Bera p-value	> 0.10	108	17	15.74
Min Spatial Autocorrelation p-value	> 0.10	20	0	0

Table 13. Summary of variable significance from the exploratory regression analysis of the elevation derivatives for the June 8, 2015 MSAVI2 values.

Variable	% Significant	% Negative	% Positive
ELEVATION	100	100	0
ASPECT	100	100	0
DIRECTRADIATION	100	75.56	24.44
SOLARRADIATION	100	95.56	4.44
DIFFUSERADIATION	95.83	81.25	18.75
DIRECTDURATION	94.34	96.23	3.77
SLOPE	92.59	77.78	22.22

Table 14. Summary of multicollinearity from the exploratory regression analysis of the elevation derivatives for the June 8, 2015 MSAVI2 values.

Variable	VIF	Violations	Covariates
ELEVATION	2.78	0	--------
ASPECT	1.31	0	--------
DIRECTRADIATION	995.88	13	SOLARRADIATION (39.39)
DIRECTDURATION	5.81	0	--------
DIFFUSERADIATION	10.59	6	--------
SOLARRADIATION	978.24	13	DIRECTRADIATION (39.39)
SLOPE	3.89	0	--------

Table 15. Summary of residual normality (JB) from the exploratory regression analysis of the elevation derivatives for the June 8, 2015 MSAVI2 values.

JB	AdjR2	AICc	K(BP)	VIF	SA	Model
0.889862	0.014703	-30967.85502	0	1	0	-ASPECT***
0.676807	0.016711	-30982.05054	0	1.022442	0	-ASPECT*** +SLOPE***
0.331885	0.073205	-31421.68688	0	2.189112	0	-ELEVATION*** -DIRECTDURATION*** -SLOPE***

Table 16. Summary of residual spatial autocorrelation (SA) from the exploratory regression analysis of the elevation derivatives for the June 8, 2015 MSAVI2 values.

SA	AdjR2	AICc	JB	K(BP)	VIF	Model
0	0.131212	-31900.0044	0	0	10.593599	-ELEVATION*** -ASPECT*** -DIRECTDURATION*** -DIFFUSERADIATION -SOLARRADIATION*** -SLOPE***
0	0.131212	-31900.00427	0	0	10.373466	-ELEVATION*** -ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION** -SLOPE***
0	0.13115	-31900.47772	0	0	3.121726	-ELEVATION*** -ASPECT*** -DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***

The results of the exploratory regression analysis are shown in tables 17 – 22 for the July 12, 2015 MSAVI2 values against the elevation and topography derivative explanatory variables (elevation, slope, aspect, total incoming solar radiation, diffuse solar radiation, direct solar radiation, and the duration of direct solar radiation). Only the top 3 regression results (highest adjusted R-squared result) for each possible number of combinations are included. Results were severe data multicollinearity or data redundancy occurred were not included.

Table 17. Highest adjusted R-squared results from the exploratory regression analysis of the elevation derivatives for the July 12, 2015 MSAVI2 values.

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.01	-20074.56	0	0	1	0	-DIFFUSERADIATION***
0.01	-20074.18	0	0	1	0	-DIRECTDURATION***
0.01	-20064.09	0	0.13	1	0	-ELEVATION***
0.02	-20134.77	0	0	1.09	0	-DIRECTRADIATION*** -DIRECTDURATION***
0.02	-20132.92	0	0	1.05	0	-DIRECTDURATION*** -SOLARRADIATION***
0.01	-20110.37	0	0	2.32	0	-DIFFUSERADIATION*** -SLOPE***
0.02	-20198.55	0	0	2.44	0	-DIRECTRADIATION*** -DIRECTDURATION*** -SLOPE***
0.02	-20198.25	0	0	2.37	0	-DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***
0.02	-20165.04	0	0	1.26	0	-ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION***
0.03	-20213.48	0	0	2.43	0	-ASPECT*** -DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***
0.03	-20212.92	0	0	2.5	0	-ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -SLOPE***
0.02	-20196.65	0	0	3.09	0	+ELEVATION -DIRECTRADIATION*** -DIRECTDURATION*** -SLOPE***
0.03	-20211.68	0	0	862.94	0	-ASPECT*** +DIRECTRADIATION -DIRECTDURATION*** -SOLARRADIATION -SLOPE***
0.03	-20211.68	0	0	5.62	0	-ASPECT*** -DIRECTDURATION*** -DIFFUSERADIATION -SOLARRADIATION*** -SLOPE***
0.03	-20211.68	0	0	5.54	0	-ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION -SLOPE***
0.03	-20209.96	0	0	10.59	0	+ELEVATION -ASPECT*** -DIRECTDURATION*** -DIFFUSERADIATION -SOLARRADIATION*** -SLOPE***
0.03	-20209.96	0	0	10.37	0	+ELEVATION -ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION -SLOPE***

Table 18. Percentage of search criteria passed from the exploratory regression analysis of the elevation derivatives for the July 12, 2015 MSAVI2 values.

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	108	0	0
Max Coefficient p-value	< 0.05	108	59	54.63
Max VIF Value	< 7.50	108	89	82.41
Min Jarque-Bera p-value	> 0.10	108	0	0
Min Spatial Autocorrelation p-value	> 0.10	19	0	0

Table 19. Summary of variable significance from the exploratory regression analysis of the elevation derivatives for the July 12, 2015 MSAVI2 values.

Variable	% Significant	% Negative	% Positive
ASPECT	98.15	100	0
DIRECTDURATION	96.23	100	0
DIRECTRADIATION	93.33	80	20
SOLARRADIATION	93.33	91.11	8.89
SLOPE	87.04	77.78	22.22
DIFFUSERADIATION	72.92	75	25
ELEVATION	47.17	71.7	28.3

Table 20. Summary of multicollinearity from the exploratory regression analysis of the elevation derivatives for the July 12, 2015 MSAVI2 values.

Variable	VIF	Violations	Covariates
ELEVATION	2.78	0	--------
ASPECT	1.31	0	--------
DIRECTRADIATION	995.88	13	SOLARRADIATION (39.39)
DIRECTDURATION	5.81	0	--------
DIFFUSERADIATION	10.59	6	--------
SOLARRADIATION	978.24	13	DIRECTRADIATION (39.39)
SLOPE	3.89	0	--------

Table 21. Summary of residual normality (JB) from the exploratory regression analysis of the elevation derivatives for the July 12, 2015 MSAVI2 values.

JB	AdjR2	AICc	K(BP)	VIF	SA	Model
0	0.003684	-20044.91173	0.167552	1	0	-DIRECTRADIATION***
0	0.001083	-20025.49882	0.521768	1	0	-ASPECT***
0	0.006247	-20064.09363	0.125851	1	0	-ELEVATION***

Table 22. Summary of residual spatial autocorrelation (SA) from the exploratory regression analysis of the elevation derivatives for the July 12, 2015 MSAVI2 values.

SA	AdjR2	AICc	JB	K(BP)	VIF	Model
0	0.026376	-20213.48264	0	0	2.428162	-ASPECT*** -DIRECTDURATION*** -SOLARRADIATION*** -SLOPE***
0	0.026303	-20212.92438	0	0	2.49793	-ASPECT*** -DIRECTRADIATION*** -DIRECTDURATION*** -SLOPE***
0	0.026271	-20211.6767	0	0	862.93929	-ASPECT*** +DIRECTRADIATION -DIRECTDURATION*** -SOLARRADIATION -SLOPE***

The results of the exploratory regression analysis are shown in tables 23 – 28 for the August 25, 2015 MSAVI2 values against the elevation and topography derivative explanatory variables (elevation, slope, aspect, total incoming solar radiation, diffuse solar radiation, direct solar radiation, and the duration of direct solar radiation). Only the top 3 regression results (highest adjusted R-squared result) for each possible number of combinations are included. Results were severe data multicollinearity or data redundancy occurred were not included.

Table 23. Highest adjusted R-squared results from the exploratory regression analysis of the elevation derivatives for the August 25, 2015 MSAVI2 values.

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.07	-15538.29	0	0	1	0	-DIRECTDURATION***
0.07	-15520.8	0	0	1	0	-DIFFUSERADIATION***
0.03	-15178.75	0	0	1	0	+SLOPE***
0.1	-15790.13	0	0	1.49	0	+ELEVATION*** -DIFFUSERADIATION***
0.08	-15591.16	0	0	2.77	0	-DIRECTDURATION*** -DIFFUSERADIATION***
0.08	-15589.86	0	0	1.05	0	-DIRECTDURATION*** -SOLARRADIATION***
0.15	-16212.07	0	0	5.67	0	+ELEVATION*** -DIFFUSERADIATION*** -SLOPE***
0.1	-15792.04	0	0	235.59	0	+ELEVATION*** +DIRECTRADIATION*** -SOLARRADIATION***
0.1	-15792.04	0	0	1.63	0	+ELEVATION*** -DIFFUSERADIATION*** -SOLARRADIATION*
0.15	-16220.62	0	0	899.39	0	+ELEVATION*** +DIRECTRADIATION*** -SOLARRADIATION*** -SLOPE***
0.15	-16220.62	0	0	5.68	0	+ELEVATION*** -DIFFUSERADIATION*** -SOLARRADIATION*** -SLOPE***
0.15	-16220.62	0	0	5.71	0	+ELEVATION*** -DIRECTRADIATION*** -DIFFUSERADIATION*** -SLOPE***
0.16	-16249.77	0	0	10.06	0	+ELEVATION*** -DIRECTDURATION*** -DIFFUSERADIATION*** -SOLARRADIATION*** -SLOPE***
0.16	-16249.77	0	0	9.82	0	+ELEVATION*** -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION*** -SLOPE***
0.15	-16219	0	0	901.65	0	+ELEVATION*** -ASPECT +DIRECTRADIATION*** -SOLARRADIATION*** -SLOPE***
0.16	-16248.9	0	0	10.59	0	+ELEVATION*** +ASPECT -DIRECTDURATION*** -DIFFUSERADIATION*** -SOLARRADIATION*** -SLOPE***
0.16	-16248.9	0	0	10.37	0	+ELEVATION*** +ASPECT -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION*** -SLOPE***

Table 24. Percentage of search criteria passed from the exploratory regression analysis of the elevation derivatives for the August 25, 2015 MSAVI2 values.

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	108	0	0
Max Coefficient p-value	< 0.05	108	60	55.56
Max VIF Value	< 7.50	108	89	82.41
Min Jarque-Bera p-value	> 0.10	108	0	0
Min Spatial Autocorrelation p-value	> 0.10	20	0	0

Table 25. Summary of variable significance from the exploratory regression analysis of the elevation derivatives for the August 25, 2015 MSAVI2 values.

Variable	% Significant	% Negative	% Positive
DIFFUSERADIATION	100	100	0
SLOPE	100	77.78	22.22
DIRECTDURATION	96.23	100	0
DIRECTRADIATION	80	46.67	53.33
SOLARRADIATION	80	91.11	8.89
ELEVATION	77.36	3.77	96.23
ASPECT	44.44	31.48	68.52

Table 26. Summary of multicollinearity from the exploratory regression analysis of the elevation derivatives for the August 25, 2015 MSAVI2 values.

Variable	VIF	Violations	Covariates
ELEVATION	2.78	0	--------
ASPECT	1.31	0	--------
DIRECTRADIATION	995.88	13	SOLARRADIATION (39.39)
DIRECTDURATION	5.81	0	--------
DIFFUSERADIATION	10.59	6	--------
SOLARRADIATION	978.24	13	DIRECTRADIATION (39.39)
SLOPE	3.89	0	--------

Table 27. Summary of residual normality (JB) from the exploratory regression analysis of the elevation derivatives for the August 25, 2015 MSAVI2 values.

JB	AdjR2	AICc	K(BP)	VIF	SA	Model
0	-0.00013	-14984.03913	0	1	0	+DIRECTRADIATION
0	0.000971	-14992.2664	0	1	0	+ASPECT***
0	-0.00013	-14984.04816	0	1	0	-ELEVATION

Table 28. Summary of residual spatial autocorrelation (SA) from the exploratory regression analysis of the elevation derivatives for the August 25, 2015 MSAVI2 values.

SA	AdjR2	AICc	JB	K(BP)	VIF	Model
0	0.156662	-16248.89767	0	0	10.593599	+ELEVATION*** +ASPECT -DIRECTDURATION*** -DIFFUSERADIATION*** -SOLARRADIATION*** -SLOPE***
0	0.156662	-16248.89763	0	0	10.373466	+ELEVATION*** +ASPECT -DIRECTRADIATION*** -DIRECTDURATION*** -DIFFUSERADIATION*** -SLOPE***
0	0.156648	-16249.77448	0	0	10.057151	+ELEVATION*** -DIRECTDURATION*** -DIFFUSERADIATION*** -SOLARRADIATION*** -SLOPE***

3.3.4 Exploratory regression – Soil Electrical Conductivity Data

The results of the exploratory regression analysis are shown in tables 29 – 34 for the June 8, 2015 MSAVI2 values against the soil EC explanatory variables (0.5 meter and 1 meter depths). Only the top 3 regression results (highest adjusted R-squared result) for each possible number of combinations are included. Severe data multicollinearity or data redundancy did not occur in this analysis.

Table 29. Highest adjusted R-squared results from the exploratory regression analysis of the soil EC data for the June 8, 2015 MSAVI2 values.

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.02	-20742.52	0.06	0.67	1	0	-CONDUCT_1***
0.01	-20676.14	0.1	0.02	1	0	-CONDUCT_05***
0.03	-20769.31	0.05	0.27	4.27	0	+CONDUCT_05*** -CONDUCT_1***

Table 30. Percentage of search criteria passed from the exploratory regression analysis of the soil EC data for the June 8, 2015 MSAVI2 values.

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	3	0	0
Max Coefficient p-value	< 0.05	3	3	100
Max VIF Value	< 7.50	3	3	100
Min Jarque-Bera p-value	> 0.10	3	0	0
Min Spatial Autocorrelation p-value	> 0.10	3	0	0

Table 31. Summary of variable significance from the exploratory regression analysis of the soil EC data for the June 8, 2015 MSAVI2 values.

Variable	% Significant	% Negative	% Positive
CONDUCT_05	100	50	50
CONDUCT_1	100	100	0

Table 32. Summary of multicollinearity from the exploratory regression analysis of the soil EC data for the June 8, 2015 MSAVI2 values.

Variable	VIF	Violations	Covariates
CONDUCT_05	4.27	0	--------
CONDUCT_1	4.27	0	--------

Table 33. Summary of residual normality (JB) from the exploratory regression analysis of the soil EC data for the June 8, 2015 MSAVI2 values.

JB	AdjR2	AICc	K(BP)	VIF	SA	Model
0.095782	0.008248	-20676.13964	0.022488	1	0	-CONDUCT_05***
0.063999	0.02118	-20742.51668	0.67436	1	0	-CONDUCT_1***
0.045333	0.026546	-20769.31248	0.273782	4.273193	0	+CONDUCT_05*** -CONDUCT_1***

Table 34. Summary of residual spatial autocorrelation (SA) from the exploratory regression analysis of the soil EC data for the June 8, 2015 MSAVI2 values.

SA	AdjR2	AICc	JB	K(BP)	VIF	Model
0	0.02118	-20742.51668	0.063999	0.67436	1	-CONDUCT_1***
0	0.026546	-20769.31248	0.045333	0.273782	4.273193	+CONDUCT_05*** -CONDUCT_1***
0	0.008248	-20676.13964	0.095782	0.022488	1	-CONDUCT_05***

The results of the exploratory regression analysis are shown in tables 35 – 40 for the July 12, 2015 MSAVI2 values against the soil EC explanatory variables (0.5 meter and 1 meter depths). Only the top 3 regression results (highest adjusted R-squared result) for each possible number of combinations are included. Severe data multicollinearity or data redundancy did not occur in this analysis.

Table 35. Highest adjusted R-squared results from the exploratory regression analysis of the soil EC data for the July 12, 2015 MSAVI2 values.

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.04	-14004.78	0	0	1	0	-CONDUCT_1***
0.03	-13956.39	0	0	1	0	-CONDUCT_05***
0.04	-14002.83	0	0	4.28	0	+CONDUCT_05 -CONDUCT_1***

Table 36. Percentage of search criteria passed from the exploratory regression analysis of the soil EC data for the July 12, 2015 MSAVI2 values.

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	3	0	0
Max Coefficient p-value	< 0.05	3	2	66.67
Max VIF Value	< 7.50	3	3	100
Min Jarque-Bera p-value	> 0.10	3	0	0
Min Spatial Autocorrelation p-value	> 0.10	3	0	0

Table 37. Summary of variable significance from the exploratory regression analysis of the soil EC data for the July 12, 2015 MSAVI2 values.

Variable	% Significant	% Negative	% Positive
CONDUCT_1	100	100	0
CONDUCT_05	50	50	50

Table 38. Summary of multicollinearity from the exploratory regression analysis of the soil EC data for the July 12, 2015 MSAVI2 values.

Variable	VIF	Violations	Covariates
CONDUCT_05	4.28	0	--------
CONDUCT_1	4.28	0	--------

Table 39. Summary of residual normality (JB) from the exploratory regression analysis of the soil EC data for the July 12, 2015 MSAVI2 values.

JB	AdjR2	AICc	K(BP)	VIF	SA	Model
0	0.037058	-14002.82773	0	4.277894	0	+CONDUCT_05 -CONDUCT_1***
0	0.037239	-14004.77989	0	1	0	-CONDUCT_1***
0	0.02798	-13956.38585	0	1	0	-CONDUCT_05***

Table 40. Summary of residual spatial autocorrelation (SA) from the exploratory regression analysis of the soil EC data for the July 12, 2015 MSAVI2 values.

SA	AdjR2	AICc	JB	K(BP)	VIF	Model
0	0.037239	-14004.77989	0	0	1	-CONDUCT_1***
0	0.037058	-14002.82773	0	0	4.277894	+CONDUCT_05 -CONDUCT_1***
0	0.02798	-13956.38585	0	0	1	-CONDUCT_05***

The results of the exploratory regression analysis are shown in tables 41 – 46 below for the August 25, 2015 MSAVI2 values against the soil EC explanatory variables (0.5 meter and 1 meter depths). Only the top 3 regression results (highest adjusted R-squared result) for each possible number of combinations are included. Severe data multicollinearity or data redundancy did not occur in this analysis.

Table 41. Highest adjusted R-squared results from the exploratory regression analysis of the soil EC data for the August 25, 2015 MSAVI2 values.

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.02	-9380.98	0	0	1	0	-CONDUCT_05***
0.01	-9339.62	0	0.1	1	0	-CONDUCT_1***
0.02	-9390.84	0	0	4.25	0	-CONDUCT_05*** +CONDUCT_1**

Table 42. Percentage of search criteria passed from the exploratory regression analysis of the soil EC data for the August 25, 2015 MSAVI2 values.

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	3	0	0
Max Coefficient p-value	< 0.05	3	3	100
Max VIF Value	< 7.50	3	3	100
Min Jarque-Bera p-value	> 0.10	3	0	0
Min Spatial Autocorrelation p-value	> 0.10	3	0	0

Table 43. Summary of variable significance from the exploratory regression analysis of the soil EC data for the August 25, 2015 MSAVI2 values.

Variable	% Significant	% Negative	% Positive
CONDUCT_1	100	100	0
CONDUCT_05	100	50	50

Table 44. Summary of multicollinearity from the exploratory regression analysis of the soil EC data for the August 25, 2015 MSAVI2 values.

Variable	VIF	Violations	Covariates
CONDUCT_05	4.25	0	--------
CONDUCT_1	4.25	0	--------

Table 45. Summary of residual normality (JB) from the exploratory regression analysis of the soil EC data for the August 25, 2015 MSAVI2 values.

JB	AdjR2	AICc	K(BP)	VIF	SA	Model
0	0.017209	-9390.842987	0.000001	4.247575	0	-CONDUCT_05*** +CONDUCT_1**
0	0.006992	-9339.618217	0.100234	1	0	-CONDUCT_1***
0	0.015092	-9380.980299	0	1	0	-CONDUCT_05***

Table 46. Summary of residual spatial autocorrelation (SA) from the exploratory regression analysis of the soil EC data for the August 25, 2015 MSAVI2 values.

SA	AdjR2	AICc	JB	K(BP)	VIF	Model
0	0.006992	-9339.618217	0	0.100234	1	-CONDUCT_1***
0	0.017209	-9390.842987	0	0.000001	4.247575	-CONDUCT_05*** +CONDUCT_1**
0	0.015092	-9380.980299	0	0	1	-CONDUCT_05***

3.4 Discussion

The results of this study outlined in this chapter indicate that the 2015 potato crop canopy health as measured by MSAVI2 analysis of high spatial resolution satellite imagery in season does vary within the field and temporally as the season progresses. The field mean MSAVI2 value started at 0.258325 on June 8, 2015 at the early vegetative stage and increased to 0.898117 by July 12, 2015 at the tuber bulking stage. The field mean MSAVI2 values drop to 0.545860 by the August 25, 2015 date during the last few days of maturation just prior to vine kill. The MSAVI2 field standard deviation values

shows variability consistently increasing during the season indicating increased variance in crop health as the season progressed.

The results of the analysis of complete spatial randomness show that the spatial patterns of the MSAVI2 feature class points and corresponding MSAVI2 values are not random in nature. This is not surprising given that the agricultural field and agronomic steps practiced year over year on the field site, such as tillage, fertility, planting, crop protection, and irrigation continually lead to crop uniformity within the field. With that said, it is logical that spatial variability within the crop canopy could be the result of other factors that can more strongly influence plant health than the standard agronomic practices such as significant differences in soil texture, topography, combinations of both, or other factors and inherent combinations. Also, the agronomic practices are not varied across the field meaning applications and practice may not be sufficiently calibrated to meet the specific needs of every area within the field that has underlying variability significantly driving crop canopy health and biomass during the season.

The exploratory regression analysis of the explanatory elevation derivatives and soil EC values and 0.5 and 1 meter depths shows that much work still needs to be done to understand what is driving the spatial crop canopy health patterns. The fact that these patterns are not random suggests some type of explanatory variable or combination of explanatory variables is driving the spatial canopy health patterns but soil EC, elevation, slope, aspect, and solar radiation derivatives do not sufficiently explain the crop canopy

patterns at this time. However, given that only three image dates were available for the growing season could indicate that better imagery frequency is needed to better ascertain the root causes causing crop heterogeneity. Also, other factors that are difficult or impossible to measure could also be at work including micro weather patterns within the field, pest and/or disease outbreaks, and man-made management decisions or equipment factors that lead to in-field variances that influence crop inputs such as fertilizer, seed, and irrigation water applications.

3.5 Conclusion

In conclusion, this research shows crop health variability measured by remote sensing methodologies does exist in the agricultural field study area. It also shows that the variability is not random but driven by other factor. However, typical differences in elevation, topographical features, incoming solar energy, and soil variability as measured by electrical conductivity mapping are not strongly related to the variability. It is likely other factors relating to inefficiencies in the irrigation system, fertility application program, and other management and cultural practices are strongly influencing spatial crop health patterns.

4 Spectral Data Collection and Analysis of PVY-Infected and Non-Infected Potato Plants

4.1 Introduction

PVY is one of the most important disease impacting the global potato industry and can be responsible for yield losses of up to 80% to 90% (Kerlan 2008; Nolte et al. 2004). There is no known cure to eradicate PVY infection in potatoes and global potato production relies on PVY-free seed to maximize yield and quality (Strand and Rude 2006). However, the lack of an accurate field-level, in-season detection system coupled with the advent of new potato varieties and viral strains making foliar symptoms less apparent to the human eye continues to aggravate disease mitigation practices (Nolte et al. 2004; Whitworth, Nolte, et al. 2010).

Previous research has shown that remotely sensed data can be used to accurately identify plant foliar plant disease symptoms caused by fungal or viral pathogens. For example, VIs where shown to be effective in identifying sugarbeet plants infected with Rhizoctonia crown and root rot at the onset of initial foliar wilting symptoms (Reynolds et al. 2012). Potato Yellow Vein Virus has also been detected using remote sensing methodologies before symptoms of chlorosis were apparent to the trained eye (P. Chávez et al. 2009).

PVY is a single-strand RNA potyvirus that consists of multiple strains including PVY^{O}, PVY^{N}, PVY^{C}, and PVY^{NTN} (Abdelhaq et al. 2006). It infects multiple plant

species resulting in a varying degree of foliar symptoms including leaf wrinkling, chlorosis, and necrotic leaf spots (Kogovšek et al. 2011). All PVY strains can cause internal and/or external tuber necrosis but the severity depends upon the potato variety and virus strain (Baarlen et al. 2005). Research has shown that viral cytoplasmic inclusion bodies, the protein structures produced by the expression of the virus genome, from necrotic and ordinary PVY strains are present in epidermal, mesophyll and vascular tissues of infected potato plants which may aid in the movement of the virus from cell to cell which aids in the overall propagation of the virus (Otulak and Garbaczewska 2012). Studies have also shown that virus transmission occurs via physical penetration of the cell wall typically by insect vectors and virus particles accumulate in leaf and stem tissues of potato plants as well as in shoot tips, roots, and tubers with the highest amount of viral RNA and cytoplasmic inclusion bodies accumulating in epidermis and trichome cells (Kogovšek et al. 2011). Reactive oxygen molecules, including hydrogen peroxide, have been shown to be part of the potato plant immune response after infection by PVY with the most intensive accumulation in the bordering cell walls of necrotic mesophyll cells and adjacent non-necrotic mesophyll cells resulting in necrotic lesions (Otulak and Garbaczewska 2010). Upon the infection, PVY has been linked to multiple gene expression response modifications in as little as 14 days after initial infection including stress-related genes, heat shock protein expression, wound inducing genes, and those involved in photosynthesis pathways (Pompe-Novak et al. 2006).

The goal of this project is to assess the spectral reflectance patterns and produce spectral signature/s for the mapping and detection of PVY-infected plants in an agricultural production field in southeastern Idaho. It is expected that the analysis will show differences in how light reflects from PVY-infected and non-infect potato plants and that a diagnostic tool can be developed to better identify PVY infected plants in commercial seed grower fields to help eradicate the disease from the industry seed pipeline.

Chapter 4 explores the following research questions:

1. Based on spectral data collected with a field spectrometer, do spectral differences exist between PVY-infected plants and nearby virus-free plants?
2. Can a spectral signature be developed to accurately differentiate PVY-infected plants within a production field?

4.2 Materials and Methods

The application of more advanced data and analytics including hyperspectral data and multifractal and spectral angle mapper analysis show much promise to identify disease incidence at the plant level (Perla Chávez et al. 2010; Yang and Everitt 2012). Support Vector Machine (SVM), a supervised classification that often shows good results in complex or noisy data will also be used. Recent research indicates very high accuracies

when classifying various plant diseases including those induced by fungal, viral and bacterial pathogens (D.Pujari et al. 2016).

4.2.1 Field Site

The field site encompasses 150 acres and is located in southeastern Idaho at an altitude of about 1,285 meters. The field was chosen because of its crop rotation, standard irrigation system, and typical agronomic practices. Coordinates are not being provided because the grower has requested the field location remain anonymous. The crop heterogeneity and field variability were analyzed in Chapter 3 and show that the potato crop was heterogeneous as measured by remotes sensing derived VIs and the patters were not completely spatially random.

The field site is managed with a typical potato and cereal grain rotation of one year planted to potatoes followed by two years planted to cereal crops and then back to potatoes the following year. It is irrigated with a center pivot system equipped with a corner extension arm with water supplied from an irrigation well supplied by the Eastern Snake River Plain Aquifer. For this study, individual plants were identified in the 2015 potato crop and tested to confirm the presence of the PVY virus and to ensure adjacent control plants did not have the virus.

4.2.2 Identifying PVY-Infected Plants.

Several steps were taken in the field to find and monitor potato plants infected with PVY. After emergence when plants were approximately eight to ten inches tall, the southwest quadrant of the field was scouted multiple times to identify potential PVY-infected plants based on visual symptoms. These symptoms included stunted plant growth, wrinkly or bubbly texture on leaves, vein burning, and a slightly different green hue of the overall plant color (Figure 18).

Figure 18. The photos above show a plant infected with PVY in the field site – the left was captured earlier in the season prior to canopy row closure and the right was captured just prior to crop senescence. The early-season image shows the slightly stunted plant growth, bubbly leaf texture, and overall lighter green hue. The late-season plant shows severely stunted growth, vein burning, and necrotic lesions.

Because of the difficulty of visually identifying PVY-infected plants in the Russet Norkotah variety, it took several trips to identify enough infected plants for an appropriate study. Much of the field scouting had to be done in conditions where bright, direct sunlight was minimized, such as partly cloudy sky or late evening conditions where visual symptoms are more apparent versus under clear skies with ample direct sunlight. The plants showing visual symptoms of PVY infection were assigned a three-digit number, flagged, and two to three neighboring plants were removed in the same row (approximately 0.5 to 0.75 meters) from each side of the suspect plant to prevent healthy neighboring plants from overgrowing the weaker plant. A single petiole was pulled from each suspect plant and from a healthy control plant directly to the north approximately 3 plants away and submitted to ICIA for enzyme-linked immunosorbent assay (ELIZA) testing to confirm the presence of the disease.

In total, 31 plants were identified and confirmed to be infected with PVY in the study area within the field that spanned areas showing significant variability in crop canopy health as measured by remote sensing and studied in Chapter 3. An additional petiole from each plant that tested positive was submitted to North Dakota State University for virus strain identification using Immunocapture Multiplex Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) that is more sensitive in detecting PVY in plant sap as well as differentiating strains (Mallik et al. 2012). Coordinates for each confirmed infected plant were collected with a Trimble GeoXH handheld receiver

and a silver tinfoil pie plate was placed on the north side of each infected plant (Figure 19).

Figure 19. The image on the left shows the locations of the 31 PVY-infected potato plants (yellow dots) in the study area. Nearby plants were removed and tin foil pie plates were used to mark the infected plants.

4.2.3 Spectral data collection

As the plants were identified, spectral data was collected from the infected and neighboring healthy control plants using a FieldSpec 4 Standard-Res Spectro radiometer (Analytical Spectral Devices, a PANalytical Company) with a spectral range of 350 – 2,500 nm and a resolution of 3 nm in VNIR and 10 nm in SWIR wavelengths. Field spectroscopy is a very effective tool to measure and characterize light reflectance and can be used to support the future development of an airborne sensor (Milton et al. 2009). Spectral data was collected at a minimum of three times for later identified plants and five times for the plants identified earliest. Table 2 shows the field spectrometer

collection dates. The field spectrometer was optimized before each data collection and a white calibration was taken approximately every 20 minutes during the collection run to account for any changes in light conditions. All scans were taken of infected plants at distance of four to six inches from the sensor the infected leaf canopy to minimize any background soil reflectance effects. All efforts were made to collect data on mostly sunny days but quickly changing local weather patterns did occur interrupting some scans briefly. If partly cloudy conditions occurred during data collection, data collection was stopped until direct sunlight conditions resumed. Using the accompanying software, the spectrometer data was corrected to surface reflectance values for continued analysis. A total of 242 readings of PVY-infected plants and accompanying non-infected controls were collected for the growing season.

Table 47. Timeline of spectral data collection for each sample plant.

Sample ID	5-Jul-15	7-Jul-15	12-Jul-15	16-Jul-15	12-Aug-15
148	X		X	X	X
154	X		X	X	X
156	X		X	X	X
158	X		X	X	X
165	X		X	X	X
166	X	X	X	X	X
167	X		X	X	X
169	X		X	X	X
172	X		X	X	X
173	X		X	X	X
174	X	X	X	X	X
175	X	X	X	X	X
176	X		X	X	X
177	X		X	X	X
181	X		X	X	X
182	X		X	X	X
185	X	X	X	X	X
202		X	X	X	X
205		X	X	X	X
206		X	X	X	X
209		X	X	X	X
210		X	X	X	X
211		X	X	X	X
212		X	X	X	X
213			X	X	X
214			X	X	X
215			X	X	X
216			X	X	X
217			X	X	X
218			X	X	X
219			X	X	X

4.2.4 Unmanned Aerial Systems Data and Analysis

During the growing season, an unmanned aerial system (UAS) equipped with a multi-spectral camera collecting NIR, red, and green wavelengths was flown over the

field site on July 15, 2015 (Figure 20). The UAS utilized for the data collection was the Steadidrone Qu4D X (SteadiDrone, Industria, Knysna, Western Cape, South Africa). The unit carried a Tetracam ADC Snap (Tetracam Inc., Chatsworth, CA) multi-spectral sensor measuring green, red, and NIR wavelengths. PixelWrench2 software (Tetracam Inc., Chatsworth, CA) was used to process the data to produce the NIR, red, and green 3-band output. The tiles were stitched together using Pix4D software (Pix4D, Switzerland).

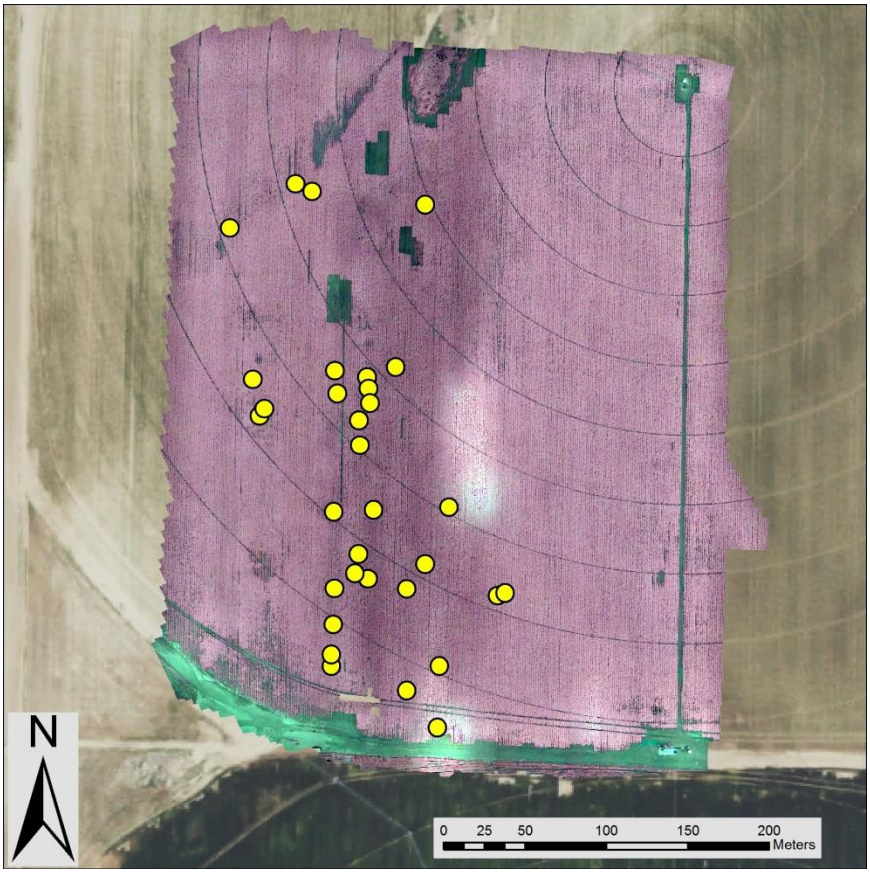

Figure 20. The figure above shows the UAS data layer as a false color layer showing the NIR band as the red channel, red as the green channel, and the green band as the blue channel overlaid with the GPS points of the PVY-infected plants. The data was capture at an approximate spatial resolution of 3 cm and includes NIR, red, and green bands.

Using the remote sensing data, canopy areas of the PVY-infected plants and of their healthy controls were digitized in ArcGIS 10.4 for Desktop to be used to extract specific pixel values to assess if a standard UAS multispectral sensor capture broad swaths of green, red, and NIR wavelengths could be used to differentiate PVY-infected plants from their nearby non-infected controls. (Figure 21). However, one plant location was not used because of a stitching error resulting in obvious pixel distortion resulting in 30 digitized canopies of PVY-infected plants and 30 digitized canopies of healthy controls. The UAS data was then processed to produce MSAVI2 values for each pixel based on the following equation (1).

$$MSAVI2 = \frac{(2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)}}{2} \quad (2)$$

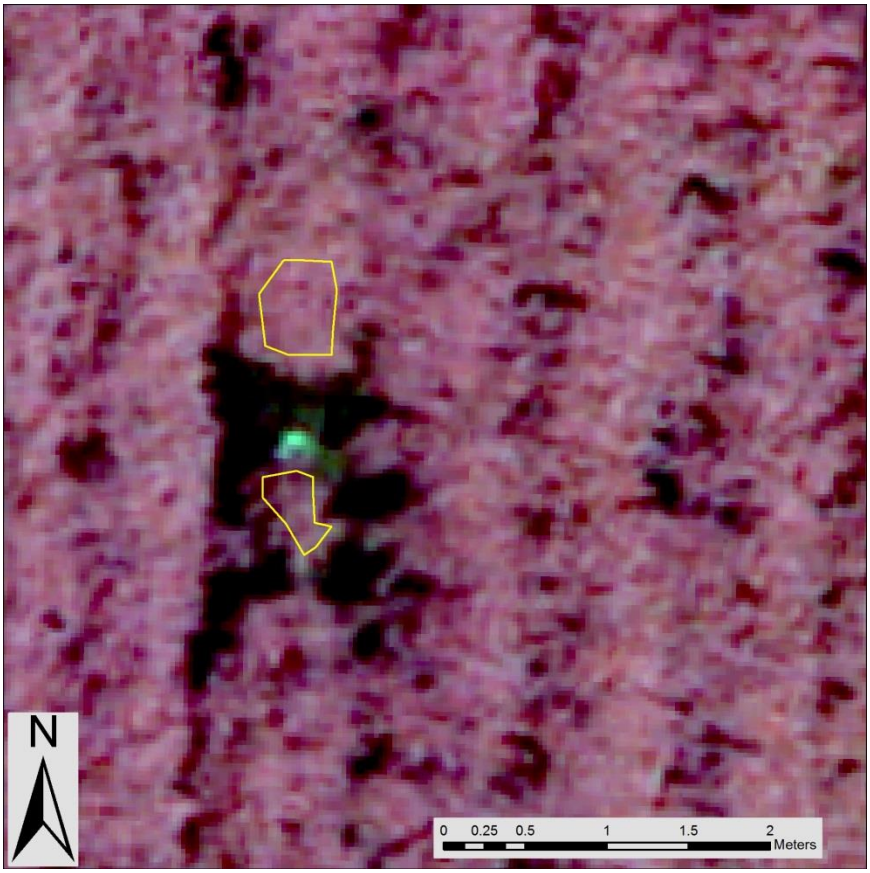

Figure 21. This image shows a section of the UAS data with two plant canopy boundaries digitized with the lower canopy being that of a PVY-infected plant and the upper one being a non-infected healthy plant. The canopy boundaries were used to extract pixel statistics for the MSAVI2 values produced from the UAS data.

Using the Zonal Statistics as Table tool in ArcGIS 10.4 for Desktop, the pixel statistics within the 60 digitized individual plant canopy geometries were extracted for the purpose of making comparisons between the PVY-infected plants and their accompanying healthy controls. The extracted MSAVI2 mean values were evaluated using the Data Analysis tools in Excel 2010 (Microsoft, Redmond, WA) first to

determine whether or not the data was normally distributed. A Normal Probability Plot was generated to compare actual values against expected values in a matching normal distribution using the Cumulative Distribution Function for the PVY-infected plants, the healthy controls, and the PVY-infected and healthy control plants as a single population. Based on those results, an F-Test was used to determine whether or not the variance of the mean MSAVI2 values of the PVY-infected plants was statistically different from the non-infected plant sample population. A t-Test was conducted to determine if the mean MSAVI2 canopy values of PVY-infected plants was statistically different from those of the virus-free population.

4.2.5 Machine learning and plant classification

As highlighted in chapter 3, the field site crop canopy consisted of significant variability that was not completely spatially random. Also shown in chapter 3, exploratory regression analysis using Ordinary Least Squares modelling showed that the variability is not related to the more commonly available elevation data and topography derivatives or moderately available soil EC data indicating crop canopy heterogeneity is driven by unknown factors at this time. Because of this, it is believed the SVM classifier is most appropriate to analyze the field spectrometer data collected from within the field site. Originally introduced as a binary classifier, SVM is a supervised classification that uses training data to fit an optimal hyper plane to classify complex data and has been

shown to be superior when separating plant species (Pouteau et al. 2012). SVM has also been shown to outperform the artificial neural network (ANN) classifier when segregating plant disease images (D.Pujari et al. 2016).

Using the Scikit-learn module in Python compiled into a script built by Dr. John Edwards, professor in the Department of Computer Science at Idaho State University, the SVM classifier was used to test whether or not the spectrometer data could be correctly differentiated into two classes: infected and non-infected. This module was selected because Python is considered to be one of the more popular programming languages for scientific computing and has an extensive amount of scientific libraries (Pedregosa et al. 2012). Specifically, C-Support Vector Classification (SVC) was used incorporating a linear kernel with the degree of the polynomial kernel function set at 3 and a range of class weights affecting the non-infected plant spectral signature training samples. The class weights were used to increase the training sample weights of the PVY-infected plant readings to reduce the number of false negative results or the type 2 error in the classification because the industry would favor that error over type 1 error or the incorrect classification of healthy plants as infected. All of the spectral signature data files collected during the growing season (242 readings) were pooled and because of the relatively small sample size, approximately 80 percent (193) of the readings were randomly selected to be used as training data. This training data included 102 spectral signature files of non-infected plants and 91 spectral signature files of PVY-infected plants leaving 49 spectral readings comprised of 19 readings of non-infected plants and

30 readings of PVY-infected plants to classify. The field spectrometer data was also subset into different regions of the electromagnetic (EM) spectrum to evaluate if certain regions delivered higher classification accuracies based on the varying weights of the training data. It was determined to evaluate the visible EM wavelengths since that is the current industry standard as well as longer wavelengths in the NIR spectrum. Table 48 shows the EM wavelengths and training data weights that were used for the classification analysis.

Table 48. SVM classification parameters.

Wavelenth Start (nm)	Wavelength End (nm)	Non-Infected Sample Weights
380	720	1, 0.8, 0.6, 0.4, 0.2
500	900	1, 0.8, 0.6, 0.4, 0.2
720	900	1, 0.8, 0.6, 0.4, 0.2
720	1300	1, 0.8, 0.6, 0.4, 0.2
900	1300	1, 0.8, 0.6, 0.4, 0.2

4.3 Results

4.3.1 PVY strain identification

Table 49 shows the results for both ICIA and NDSU laboratory testing for each plant in the study. Of the 31 plants tested in the field site, 20 were identified as $PVY^{N:O}$ (NO), 10 as the PVY^{NTN} (EU-NTN), and 1 as PVY^{O} (O). PVY^{O} is a common strain that

results in foliar symptoms ranging from mild mosaic to necrosis and was the predominant strain in North America before 1990 while tuber necrotic strains $PVY^{N:O}$ and PVY^{NTN} have become more common in recent years in North America and are more difficult to identify based on visual foliar symptoms (Halterman et al. 2012; Nolte et al. 2002).

Table 49. ICIA and NDSU laboratory testing results for the presence of PVY and virus strain for each sample plant.

Sample ID	ICIA Confirmation	NDSU Strain
148	yes	EU-NTN
154	yes	NO
156	yes	NO
158	yes	EU-NTN
165	yes	EU-NTN
166	yes	O
167	yes	NO
169	yes	NO
172	yes	NO
173	yes	NO
174	yes	NO
175	yes	NO
176	yes	NO
177	yes	EU-NTN
181	yes	NO
182	yes	NO
185	yes	NO
202	yes	EU-NTN
205	yes	NO
206	yes	NO
209	yes	NO
210	yes	NO
211	yes	NO
212	yes	EU-NTN
213	yes	NO
214	yes	NO
215	yes	EU-NTN
216	yes	EU-NTN
217	yes	NO
218	yes	EU-NTN
219	yes	EU-NTN

4.3.2 Unmanned Aerial Systems Data and Analysis Results

The results of the mean MSAVI2 values derived from the UAS multi-spectral data captured July 15, 2015 and extracted from the digitized plant canopy polygons are shown in table 50. The mean MSAVI2 value for PVY-infected plant canopies is 0.7597 and the mean MSAVI2 value for the non-infected plants is 0.763. The average difference between all infected and non-infected plants is 0.003 or a percent difference of 0.453 percent. Also, 11 infected plants have a higher MSAVI2 value than the neighboring non-infected control (Figure 22).

Table 50. Plant canopy MSAVI2 statistics from UAS derived multispectral data collected July 15, 2015.

Sample ID	PVY MSAVI2 mean values	Control MSAVI2 mean values	PVY MSAVI2 std. values	Control MSAVI2 std. values	PVY MSAVI2 var. values	Control MSAVI2 var. values
148	0.745728172	0.742020091	0.033652073	0.024030108	0.001132462	0.000577446
154	0.771771267	0.770007293	0.02889973	0.023373778	0.000835194	0.000546334
156	0.766594751	0.767098079	0.031397013	0.027746175	0.000985772	0.00076985
158	0.792007655	0.780200622	0.044702704	0.041125348	0.001998332	0.001691294
165	0.74418363	0.760980692	0.030931685	0.023550182	0.000956769	0.000554611
166	0.774123633	0.775835488	0.035783921	0.029093886	0.001280489	0.000846454
167	0.799500084	0.796913014	0.051521567	0.021902402	0.002654472	0.000479715
169	0.776466458	0.770240871	0.042297422	0.027556508	0.001789072	0.000759361
172	0.769320282	0.7675479	0.035701037	0.021330331	0.001274564	0.000454983
173	0.751806684	0.774113436	0.036896048	0.023824716	0.001361318	0.000567617
174	0.782744916	0.760286164	0.038943901	0.023158384	0.001516627	0.000536311
175	0.751529659	0.7624641	0.027191575	0.02858325	0.000739382	0.000817002
176	0.755050642	0.77969146	0.039523051	0.026923948	0.001562072	0.000724899
177	0.733410577	0.752160307	0.026736482	0.023826713	0.000714839	0.000567712
181	0.720857095	0.746388022	0.039638418	0.014442591	0.001571204	0.000208588
182	0.74125977	0.744665102	0.024142319	0.021097469	0.000582852	0.000445103
185	0.787474637	0.788327601	0.029585669	0.02253591	0.000875312	0.000507867
205	0.710701441	0.713502003	0.050954012	0.028794781	0.002596311	0.000829139
206	0.768825546	0.749482609	0.031551828	0.032991679	0.000995518	0.001088451
209	0.765234098	0.77112358	0.031772567	0.027185017	0.001009496	0.000739025
210	0.758432928	0.763259278	0.029973442	0.023957932	0.000898407	0.000573982
211	0.752195033	0.77674762	0.042960871	0.026372391	0.001845636	0.000695503
212	0.767331968	0.760250214	0.031727688	0.035620824	0.001006646	0.001268843
213	0.76383807	0.768365394	0.015162181	0.015284986	0.000229892	0.000233631
214	0.757860431	0.763756933	0.030713696	0.023802048	0.000943331	0.000566537
215	0.774824182	0.776197007	0.035943627	0.029709846	0.001291944	0.000882675
216	0.762226842	0.759459147	0.044358023	0.028730585	0.001967634	0.000825446
217	0.754071554	0.75966059	0.026906216	0.024361811	0.000723944	0.000593498
218	0.754376627	0.74602407	0.028467454	0.02398827	0.000810396	0.000575437
219	0.737244161	0.743607019	0.031637898	0.018845158	0.001000957	0.00035514

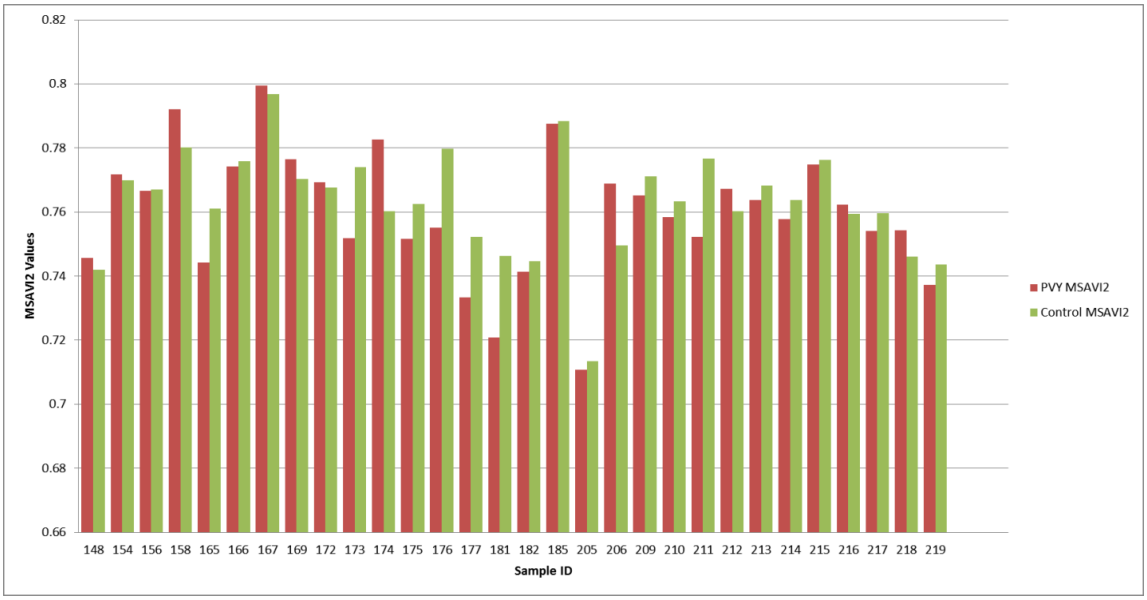

Figure 22. This graph shows the mean MSAVI2 values for each PVY-infected plant and its accompanying non-infected control neighbor. The data was derived from the July 15, 2015 multispectral data collected via UAS.

Table 51. Descriptive Statistics for the PVY-Infected, Control, and Total mean MSAVI2 values of the plant canopy polygons extracted from the UAS data collected July 15, 2015.

PVY-Infected Plants		*Control Plants*		*PVY-Infected and Control Plants*	
Mean	0.75969976	Mean	0.763012524	Mean	0.761356142
Standard Error	0.003601203	Standard Error	0.002986885	Standard Error	0.002329438
Median	0.760329885	Median	0.763508105	Median	0.762861689
Standard Deviation	0.019724598	Standard Deviation	0.016359845	Standard Deviation	0.018043752
Sample Variance	0.00038906	Sample Variance	0.000267645	Sample Variance	0.000325577
Kurtosis	0.499748916	Kurtosis	1.803181365	Kurtosis	0.847387925
Skewness	-0.357788544	Skewness	-0.69742567	Skewness	-0.520723803
Range	0.088798643	Range	0.083411011	Range	0.088798643
Minimum	0.710701441	Minimum	0.713502003	Minimum	0.710701441
Maximum	0.799500084	Maximum	0.796913014	Maximum	0.799500084
Sum	22.79099279	Sum	22.89037571	Sum	45.6813685
Count	30	Count	30	Count	60

The Normal Probability Plots for the mean PVY-infected, non-infected control, and both PVY-infected and non-infected control canopy mean MSAVI2 values extracted from the UAS data captured July 15, 2015 are shown in Figure 23. They all show that the data is closely aligned with expected values within a normally distributed pattern when compared to the standard deviation as shown by the Z-values. This indicates the data is suitable for an F-test to determine whether or not sample variances of the PVY-infected sample population and the non-infected control sample population are statistically the same.

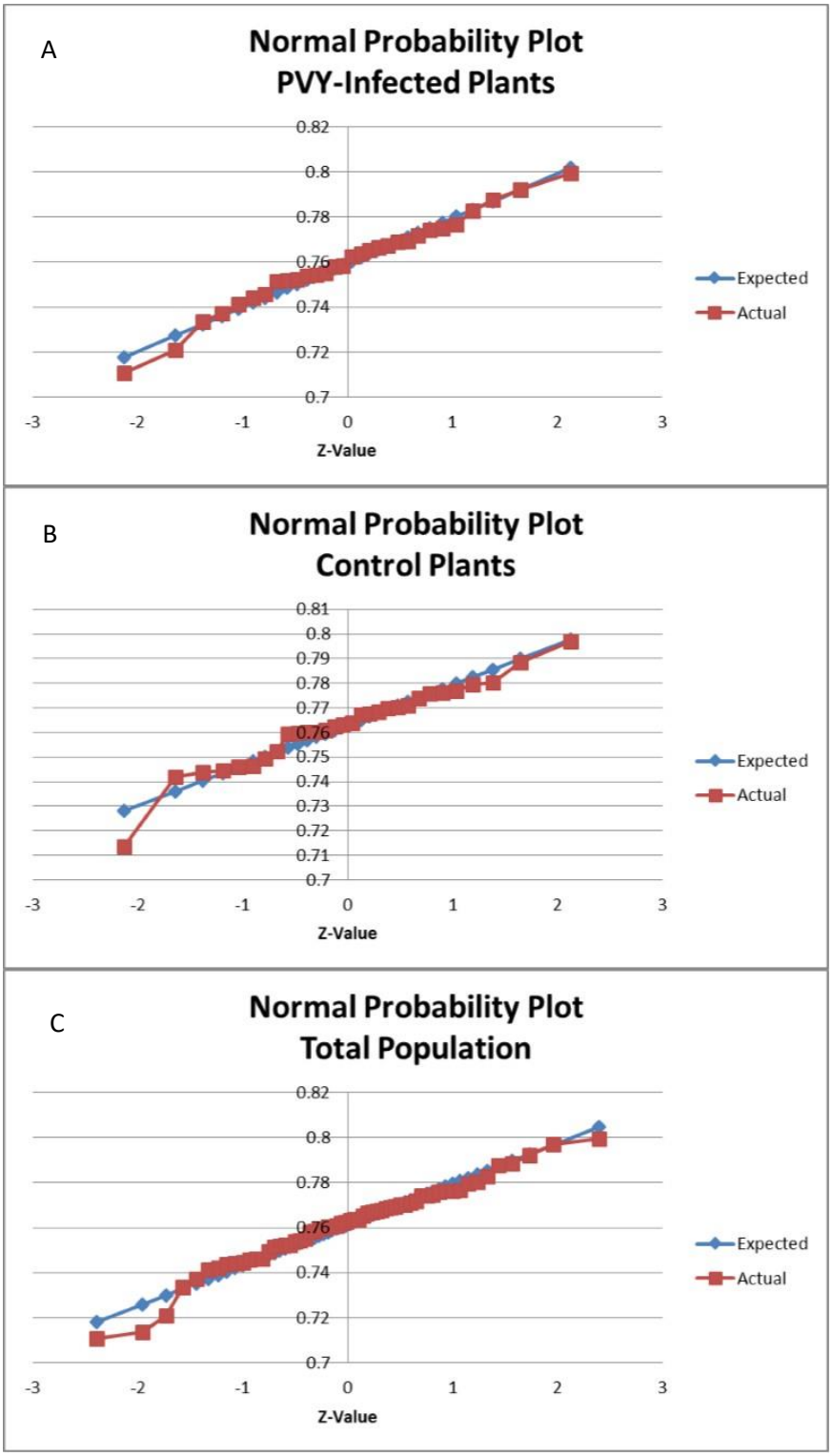
A
Normal Probability Plot
PVY-Infected Plants
0.82
0.8
0.78
0.76
0.74
0.72
0.7
-3
-2
-1
0
1
2
3
Z-Value
Expected
Actual
B
Normal Probability Plot
Control Plants
0.81
0.8
0.79
0.78
0.77
0.76
0.75
0.74
0.73
0.72
0.71
0.7
-3
-2
-1
0
1
2
3
Z-Value
Expected
Actual
C
Normal Probability Plot
Total Population
0.82
0.8
0.78
0.76
0.74
0.72
0.7
-3
-2
-1
0
1
2
3
Z-Value
Expected
Actual

Figure 23. The Normal Probability Plots for the PVY-infected (A), non-infected (B), and total population (C) mean MSAVI2 values for the extracted plant canopies digitized over the UAS data collected July 15, 2015 are shown in the graphs. The actual and expected data values are very closely aligned and indicate all of the sample populations have a normalized distribution.

The results of the F-test indicate that the null hypothesis that the mean plant canopy MSAVI2 values of the PVY-infected plants and the mean plant canopy MSAVI2 values of the non-infected plants are statistically equal at a 95% confidence interval where the probability value equals 0.16. Also, the F value of 1.453 is less than the F Critical one-tail value of 1.86 indicating that the null hypothesis cannot be rejected.

Table 52. F-test results.

	PVY MSAVI2 mean values	*Control MSAVI2 mean values*
Mean	0.75969976	0.763012524
Variance	0.00038906	0.000267645
Observations	30	30
df	29	29
F	1.453643678	
P(F<=f) one-tail	0.159694048	
F Critical one-tail	1.860811435	

The results of the t-Test for the mean plant canopy MSAVI21 values of the PVY-infected plants and the mean plant canopy MSAVI2 values of the non-infected plants indicate that the null hypothesis stating the populations are statistically equal cannot be rejected for both one- and two-tail tests.

Table 53. T-Test results.

	PVY MSAVI2 mean values	*Control MSAVI2 mean values*
Mean	0.75969976	0.763012524
Variance	0.00038906	0.000267645
Observations	30	30
Pooled Variance	0.000328352	
Hypothesized Mean Difference	0	
df	58	
t Stat	-0.708053653	
P(T<=t) one-tail	0.240873905	
t Critical one-tail	1.671552762	
P(T<=t) two-tail	0.481747811	
t Critical two-tail	2.001717484	

4.3.3 SVM Results

Table 54 shows the results of the SVM classification analysis including the classifier accuracies and the false negative (FN) rate which shows the percentage of PVY-infected plants misclassified as not-infected. The spectral readings for the various EM wavelength sections were also compiled graphically to visualize the spectral signatures (Figure 24).

Table 54. SVM classification results.

# train	# test	startWave	endWave	weight	accuracy	FN rate
193	49	380	720	1	46.94	86.7
193	49	380	720	0.8	73.47	6.7
193	49	380	720	0.6	61.22	0.0
193	49	380	720	0.4	61.22	0.0
193	49	380	720	0.2	61.22	0.0
193	49	500	900	1	89.8	13.3
193	49	500	900	0.8	89.8	13.3
193	49	500	900	0.6	83.67	13.3
193	49	500	900	0.4	77.55	13.3
193	49	500	900	0.2	75.51	6.7
193	49	720	900	1	87.76	13.3
193	49	720	900	0.8	87.76	13.3
193	49	720	900	0.6	79.59	13.3
193	49	720	900	0.4	77.55	10.0
193	49	720	900	0.2	73.47	0.0
193	49	720	1300	1	89.8	13.3
193	49	720	1300	0.8	83.67	13.3
193	49	720	1300	0.6	81.63	13.3
193	49	720	1300	0.4	81.63	10.0
193	49	720	1300	0.2	77.55	3.3
193	49	900	1300	1	87.76	13.3
193	49	900	1300	0.8	85.71	13.3
193	49	900	1300	0.6	87.76	6.7
193	49	900	1300	0.4	79.59	6.7
193	49	900	1300	0.2	73.47	3.3

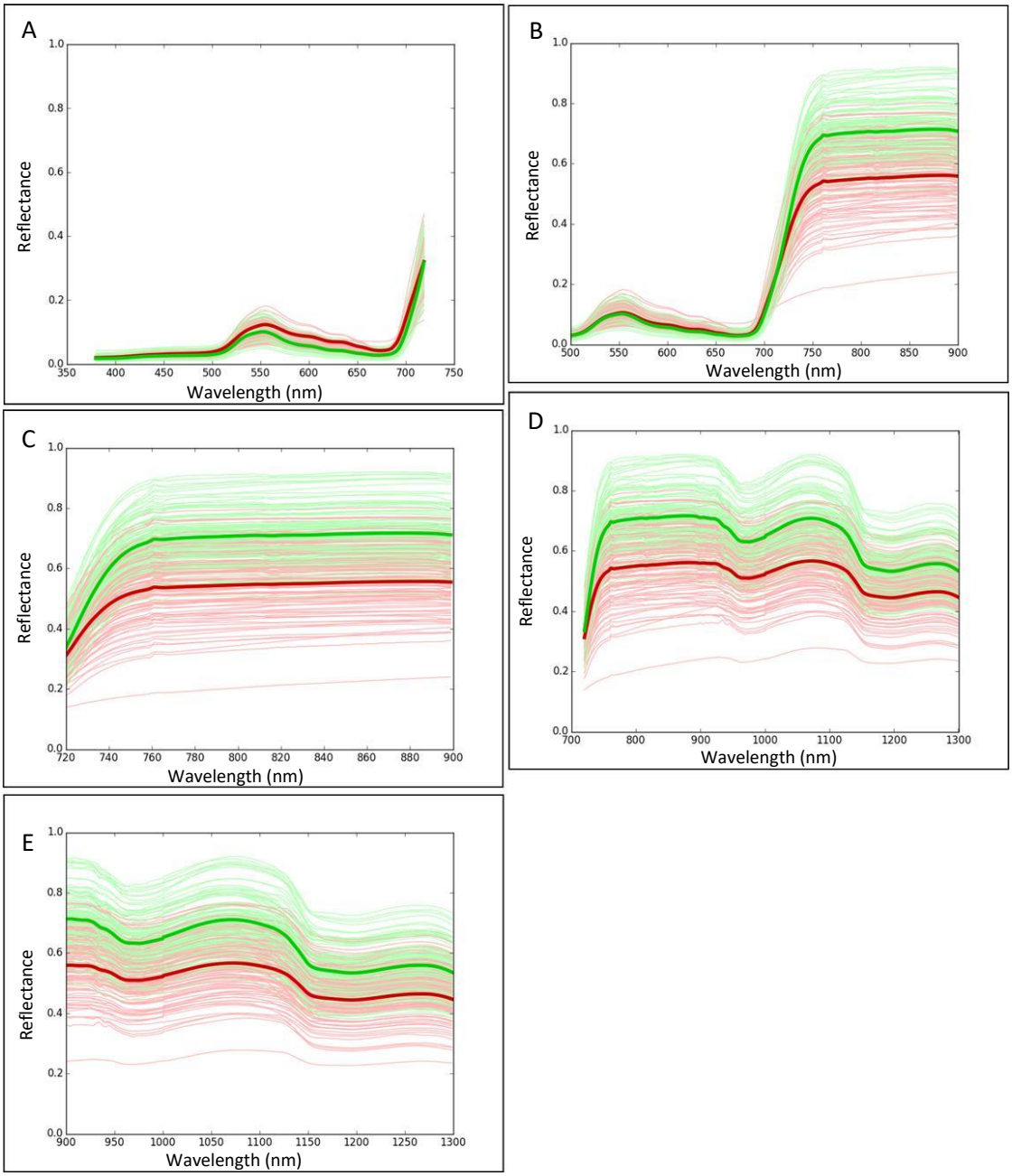

Figure 24. The graphs above show the reflectance values by wavelength of the field spectrometer data collected during the growing season and subset to specific wavelength ranges for the SVM classification analysis including 380-720 nm (A), 500-900 nm (B), 720-900 nm (C), 720-900 nm (D), and 900-1,300 nm (E). The green curves show the reflectance values for the non-infected plants and the red curves show the reflectance values for the PVY-infected plants. The prominent dark green curves show the overall average values for the healthy plants and the prominent dark red curves show the overall average reflectance values for the PVY-infected plants.

4.4 Discussion

The results of the study conducted in this chapter show that PVY was successfully identified in the field using visual field inspections and laboratory testing. However, extensive field scouting and timing with beneficial lighting conditions were incorporated to improve the accuracies as is done in the industry. The fact that of the 31 plants, 20 were identified as the $PVY^{N:O}$ strain, 10 were identified as the PVY^{NTN}, and only one identified as PVY^{O}, corroborates other research findings regarding the fact that tuber necrotic strains $PVY^{N:O}$ and PVY^{NTN} have become more common in recent years in North America and are more difficult to identify based on visual foliar symptoms versus the traditionally more common PVY^{O} strain.

The analysis of the UAS data collected July 15, 2015 indicates that 3 or 4 band multispectral sensors used to derive robust VIs such as MSAVI2 are not adequate to differentiate PVY-infected plants from non-infected plants. This approach would like be compounded by other field or agronomic variables not easily measured or incorporated into an analysis that would still affect over all plant health which could add significant bias to VI values.

The SVM classification analysis yielded very promising results, especially when incorporating NIR wavelengths not used in VIs such as NDVI or MSAVI2. The 500 nm to 900 nm range yielded the highest classification accuracy of 89.8 percent with no weight and with a weight of 0.8 applied to the non-infected plants. This was a significant

improvement over the 46.94 percent accuracy when using the 380 nm to 720 nm range with no training weight applied which is equivalent to the visual spectrum that is the current industry standard. However, applying a training weight coefficient of 0.8 to the visual wavelengths did improve the classification to 73.47 percent. The classification analysis using the 720 nm to 1,300 nm range also produced an accuracy of 89.8 percent matching that of the 500 nm to 900 nm range. The analysis of the 900 nm to 1,300 nm spectrum yielded an 87.76 percent accuracy. Overall, using the coefficients to weight the training data to favor the spectral signatures of PVY-infected plants did reduce the FN rate but also resulted in lower overall accuracies. The fact that the best accuracies of this study incorporated NIR wavelengths indicates that EM reflectance of infected leaves is revealing cellular impacts at deeper levels of the mesophyll cells within the leaf versus just the epidermal cells. This does coincide with various studies showing that the PVY virus does affect cellular structures within the mesophyll cells as well as the plant vascular system.

4.5 Conclusion

In conclusion, this research shows that PVY-infected potato plants do reflect electromagnetic energy differently when compared to non-infected plants. However, this study also shows that the current industry standard of using visual inspections is the least

accurate when compared to longer wavelengths beyond what the human eye can detect. It's also apparent from this research that standard multispectral sensors capturing large swaths of the blue, green, red, and NIR ranges are not adequate to differentiate PVY-infected plants from non-infected neighbors given the lack of statistical differences at the canopy level. It is likely a sensor or sensors capturing longer wavelengths indicative of mesophyll cell structure will produce more accurate results than current industry detection standards.

5 Conclusion

The research presented in this paper shows in-field variability does exist within this agricultural production field relative to the 2015 potato crop. It also shows that high resolution satellite imagery and likely other sources of multispectral remote sensing data of sufficient radiometric quality and adequate spatial, spectral, and temporal resolution can be used to quantify and measure the spatial patterns of canopy health. Knowing that the inherent spatial patterns of canopy healthy are not random but the result of underlying influences is a strong indication that precision agriculture practices could be successfully applied to improve overall crop production when those underlying influences can be recognized. It is also necessary to understand that the crop canopy is not homogeneous and that the variability is not completely spatially random when evaluating spectral

signatures of PVY-infected and non-infected plants located in the field to further assess how to improve classification efforts. Unfortunately, it is likely that much work and additional data is needed before the underlying influences can be quantified and mapped. The fact that it is difficult to accurately establish the factor/s causing field variability is an indication that machine learning or machine vision will be a necessity to apply this type of technology to industry, especially given the fact that factors driving field variability can be very different in various fields or regions. This research shows that PVY-infected potato plants do reflect electromagnetic energy differently when compared to non-infected plants. However, this study also shows that the current industry standard of using visual inspections is the least accurate when compared to longer wavelengths beyond what the human eye can detect.

This research also shows it is very difficult to differentiate PVY-infected plants from non-infected neighbors using traditional multispectral data processed to an MSAVI2 derivative derived from UAS sensors. It is likely that standard multispectral sensors capturing large swaths of the blue, green, red, and NIR ranges are not adequate to differentiate PVY-infected plants from non-infected neighbors given the lack of statistical differences at the canopy level based on the statistical data outlined in this paper. This is especially apparent based on the MSAVI2 statistics collected from the individual plant canopies of PVY-infected plants and their neighboring controls. Statistically, there was only a 0.433% difference between the average MSAVI2 value for PVY-infected plants and the non-infected control and 11 PVY-infected plants had higher mean MSAVI2

values which indicates using type of methodology could lead to inaccurate results and excessive type 2 error or the misclassification of sick plants as healthy. It is likely a sensor or sensors capturing longer wavelengths indicative of mesophyll cell structure will produce more accurate results than current industry detection standards based on the success of the SVM classifier applied to the field spectrometer data.

This research also shows that the industry standard of relying on human vision using only visible EM wavelengths is achieving less than 50-percent accuracy. However, it is expected that levels of training of given field inspectors could lead to significant variability in human-based classification depending upon experience, field agronomic conditions, lighting conditions, and overall plant and crop health. Although further study is needed, this research shows that PVY-infected plants could be accurately differentiated from neighboring "healthy" plants using optimal EM wavelengths and machine learning algorithms.

Much research has been done to quantify the financial impacts of PVY on grower incomes and state and regional economies. Although yield and quality metrics will vary based on variety, agronomics, market conditions, and other factors, it has been shown that Idaho potato producers collectively lose 2.3 million hundred weight (cwt) valued at approximately $14 million. Despite industry efforts of managing certified seed programs and visual identification and removal of infected plants from seed production fields, PVY continues to be a major threat to global potato production. This research shows that a new

and more accurate way of detecting and mapping PVY-infected plants could be accomplished with current remote sensing, UAS, and sensor technologies at an accuracy far superior to current industry standards. It is likely that a wide scale adoption of a remote sensing technology solutions will give seed growers the ability to remove infected plants more efficiently thereby reducing PVY inoculum and provide better seed potato quality to commercial producers to help mitigate economic harm and improve environmental sustainability by reducing dependence on pesticides currently used to control insect vectors.

6 References

Abdelhaq, H., Barker, H., Bradshaw, J. E., Coleman, W. K., Ezekiel, R., Forbes, G. A., et al. (2006). *Handbook of Potato Production, Improvement, and Postharvest Management*. (J. Gopal & S. M. P. Khurana, Eds.). The Haworth Press, Inc.

Adamchuk, V. I., Ferguson, R. B., & Hergert, G. W. (2010). Soil Heterogeneity and Crop Growth. In E. C. Oerke, R. Gerhards, G. Menz, & R. A. Sikora (Eds.), *Precision Crop Protection - the Challenge and Use of Heterogeneity* (pp. 3–16). Springer Science+Business Media.

Alvarez, J. M., Bain, P., Baird, C. R., Bechinski, E. J., Bohl, W. H., Corsini, D. L., et al. (2003). *Potato Production Systems*. (J. C. Stark & S. L. Love, Eds.). University of Idaho Agricultural Communications.

Anselin, L. (1995). Local indicators of spatial association — LISA. *Geographical Analysis*, *27*(2), 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x

Baarlen, D. ir. P. van, Bokx, D. ir. J. A. de, Brinkman, I. H., Cuperus, C., Flier, D. ir. W. G., Haar, I. H. van de, et al. (2005). *Potato diseases*. (D. ir. A. Mulder & D. ir. L. J. Turkensteen, Eds.). Aardappelwereld B.V. & NIVAP.

Bala, S. K., & Islam, a. S. (2009). Correlation between potato yield and MODIS-derived vegetation indices. *International Journal of Remote Sensing*, *30*(10), 2491–2507. doi:10.1080/01431160802552744

Boyd, N. S., Gordon, R., & Martin, R. C. (2002). Relationship between leaf area index and ground cover in potato under different management conditions. *Potato Research*, *45*, 117–129. doi:10.1007/BF02736107

C.~Cortes, C.~Cortes, V.~Vapnik, & V.~Vapnik. (1995). Support Vector Networks. *Machine Learning*, *20*(3), 273~-~297. doi:10.1007/BF00994018

Campbell, J. B. (2007). *Introduction to Remote Sensing* (Fourth.). The Guilford Press.

Chávez, P., Yarlequé, C., Piro, O., Posadas, A., Mares, V., Loayza, H., et al. (2010). Applying multifractal analysis to remotely sensed data for assessing PYVV infection in potato (Solanum tuberosum L.) crops. *Remote Sensing*, *2*(5), 1197–1216. doi:10.3390/rs2051197

Chávez, P., Zorogastúa, P., Chuquillanqui, C., Salazar, L. F., Mares, V., & Quiroz, R. (2009). Assessing Potato Yellow Vein Virus (PYVV) infection using remotely sensed data. *International Journal of Pest Management*, *55*(3), 251–256.

doi:10.1080/09670870902862685

D.Pujari, J., Yakkundimath, R., & Byadgi, A. S. (2016). SVM and ANN Based Classification of Plant Diseases Using Feature Reduction Technique. *International Journal of Interactive Multimedia and Artificial Intelligence*, *3*(7), 6. doi:10.9781/ijimai.2016.371

Dupuis, B., Schwaerzel, R., & Derron, J. (2014). Efficacy of Three Strategies Based on Insecticide , Oil and Elicitor Treatments in Controlling Aphid Populations and Potato virus Y Epidemics in Potato Fields. *Journal of Phytopathology*, *162*, 14–18. doi:10.1111/jph.12148

Getis, A., & Ord, J. K. (1992). The Analysis of Spatial Association. *Geographical Analysis*, *24*(3), 189–206. doi:10.1111/j.1538-4632.1992.tb00261.x

Gooley, L., Huang, J., Pagé, D., & Triantafilis, J. (2014). Digital soil mapping of available water content using proximal and remotely sensed data. *Soil Use and Management*, *30*(1), 139–151. doi:10.1111/sum.12094

Guo, Y., Shi, Z., Huang, J., Zhou, L., Zhou, Y., & Wang, L. (2016). Characterization of field scale soil variability using remotely and proximally sensed data and response surface method. *Stochastic Environmental Research and Risk Assessment*, *30*(3), 859–869. doi:10.1007/s00477-015-1135-0

Hadjimitsis, D. G., Papadavid, G., Agapiou, a., Themistocleous, K., Hadjimitsis, M. G., Retalis, a., et al. (2010). Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. *Natural Hazards and Earth System Science*, *10*(1984), 89–95. doi:10.5194/nhess-10-89-2010

Halterman, D., Charkowski, a, Verchot, J., & Daami-Remadi, M. (2012). Potato, viruses, and seed certification in the USA to provide healthy propagated tubers. *Pest technology*, *6*(Special Issue 1), 1–14.

Hatfield, J. L., & Prueger, J. H. (2010). Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sensing*, *2*(2), 562–578. doi:10.3390/rs2020562

Hillnhütter, C., Mahlein, A. K., Sikora, R. A., & Oerke, E. C. (2011). Remote sensing to detect plant stress induced by Heterodera schachtii and Rhizoctonia solani in sugar beet fields. *Field Crops Research*, *122*(1), 70–77. doi:10.1016/j.fcr.2011.02.007

Kerlan, C. (2008). Potato Viruses. *Encyclopedia of Virology*, 296–309. doi:DOI: 10.1016/B978-012374410-4.00646-4

Kogovšek, P., Kladnik, A., Mlakar, J., Znidarič, M. T., Dermastia, M., Ravnikar, M., & Pompe-Novak, M. (2011). Distribution of Potato virus Y in potato plant organs, tissues, and cells. *Phytopathology*, *101*(11), 1292–300. doi:10.1094/PHYTO-01-11-0020

Mahlein, A. K., Steiner, U., Dehne, H. W., & Oerke, E. C. (2010). Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precision Agriculture*, *11*(4), 413–431. doi:10.1007/s11119-010-9180-7

Mahlein, A.-K. (2016). Plant Disease Detection by Imaging Sensors - Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *Plant Disease*, *100*(2), 241–251.

Mallik, I., Anderson, N. R., & Gudmestad, N. C. (2012). Detection and Differentiation of Potato Virus Y Strains from Potato Using Immunocapture Multiplex RT-PCR. *American Journal of Potato Research*, *89*(3), 184–191. doi:10.1007/s12230-012-9241-8

McIntosh, C. S., Aryal, G. R., Watson, P., & Nolte, P. (2015). State-Wide Impacts of PVY - Idaho.

Mertens, F. M., Pätzold, S., & Welp, G. (2008). Spatial heterogeneity of soil properties and its mapping with apparent electrical conductivity. *Journal of Plant Nutrition and Soil Science*, *171*(2), 146–154. doi:10.1002/jpln.200625130

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

Moran, A. P. A. P. (2016). Biometrika Trust Notes on Continuous Stochastic Phenomena Published by : Oxford University Press on behalf of Biometrika Trust Stable URL : http://www.jstor.org/stable/2332142 Accessed : 13-04-2016 05 : 12 UTC Your use of the JSTOR archive indicates your , *37*(1), 17–23.

National Agricultural Statistics Service. (n.d.). https://quickstats.nass.usda.gov/

Nelson, M. R., Orum, T. V, Jaime-garcia, R., & Nadeem, A. (1999). Applications of geographic information systems and geostatistics in plant disease epidemiology and management. *Plant Disease*, *83*(4), 308–319.

Nolte, P., Alvarez, J. M., & Whitworth, J. L. (2002). *Potato Virus Y Management for the Seed Potato Producer*. *University of Idaho Extension*.

Nolte, P., Falls, I., Whitworth, J. L., States, U., & Thornton, M. K. (2004). Effect of Seedborne Potato virus Y on Performance of Russet Burbank , Russet Norkotah ,

and Shepody Potato. *Plant Disease*, *88*(3), 248–252.

Oliver, M. A. (2010). An Overview of Geostatistics and Precision Agriculture. In M. A. Oliver (Ed.), *Geostatistical Applications for Precision Agriculture* (pp. 1–34). Springer Science+Business Media.

Ord, J. K., & Getis, A. (2010). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, *27*(4), 286–306. doi:10.1111/j.1538-4632.1995.tb00912.x

Otulak, K., & Garbaczewska, G. (2010). Localisation of hydrogen peroxide accumulation during Solanum tuberosum cv. Rywal hypersensitive response to Potato virus Y. *Micron*, *41*(4), 327–335. doi:10.1016/j.micron.2009.12.004

Otulak, K., & Garbaczewska, G. (2012). Cytopathological Potato virus Y structures during Solanaceous plants infection. *Micron*, *43*(7), 839–850. doi:10.1016/j.micron.2012.02.015

Patzold, S., Mertens, F. M., Bornemann, L., Koleczek, B., Franke, J., Feilhauer, H., & Welp, G. (2008). Soil heterogeneity at the field scale: A challenge for precision crop protection. *Precision Agriculture*, *9*(6), 367–390. doi:10.1007/s11119-008-9077-x

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2012). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. doi:10.1007/s13398-014-0173-7.2

Peeters, A., Zude, M., K??thner, J., ??nl??, M., Kanber, R., Hetzroni, A., et al. (2015). Getis-Ord's hot- and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. *Computers and Electronics in Agriculture*, *111*, 140–150. doi:10.1016/j.compag.2014.12.011

Pompe-Novak, M., Gruden, K., Baebler, Š., Krečič-Stres, H., Kovač, M., Jongsma, M., & Ravnikar, M. (2006). Potato virus Y induced changes in the gene expression of potato (Solanum tuberosum L.). *Physiological and Molecular Plant Pathology*, *67*(3–5), 237–247. doi:10.1016/j.pmpp.2006.02.005

Pouteau, R., Meyer, J.-Y., Taputuarai, R., & Stoll, B. (2012). Support vector machines to map rare and endangered native plants in Pacific islands forests. *Ecological Informatics*, *9*, 37–46. doi:10.1016/j.ecoinf.2012.03.003

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A Modified Soil Adjusted Vegetation Index, *126*, 119–126.

Rey-Caramés, C., Diago, M. P., Pilar Martín, M., Lobo, A., & Tardaguila, J. (2015).

Using RPAS multi-spectral imagery to characterise vigour, leaf development, yield components and berry composition variability within a vineyard. *Remote Sensing*, *7*(11), 14458–14481. doi:10.3390/rs71114458

Reynolds, G. J., Windels, C. E., MacRae, I. V, & Laguette, S. (2012). Remote sensing for assessing *Rhizoctonia crown* and root rot severity in sugar beet. *Plant Disease*, *96*(4), 497–505. doi:10.1094/pdis-11-10-0831

Schmitt, D. P. (2002). Embracing the Emerging Precision Agriculture Technologies for Site-Specific Management of Yield-Limiting Factors. *Journal of Nematology*, *34*(3), 185–188.

Strand, L. L., & Rude, P. A. (2006). *Integrated Pest Management for Potatoes in the Western United States*. (M. L. Flint, Ed.) (Second.). University of California Agriculture and Natural Resources Communication Services.

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *PNAS*, *108*(50), 20260–4. doi:10.1073/pnas.1116437108

Whitworth, J., Hamm, P., & McIntosh, C. (2010). Effect of potato virus Y on yield of a clonal selection of Russet Norkotah. *American Journal of Potato Research*, *87*(3), 310–314. doi:10.1007/s12230-010-9134-7

Whitworth, J., Nolte, P., & Mcintosh, C. (2006). Effect of Potato virus Y on Yield of Three Potato Cultivars. *Plant Disease*, *90*(1), 73–76.

Whitworth, J., Nolte, P., Singh, R., Boucher, A., & Xu, H. (2010). Potato Virus Y: An evolving concern for Potato Crops in the United States and Canada. *Plant Disease*, *94*(12), 1384–1397. doi:10.1094/PDIS-02-10-0124

Wu, J., Wang, D., & Bauer, M. E. (2007). Assessing broadband vegetation indices and QuickBird data in estimating leaf area index of corn and potato canopies. *Field Crops Research*, *102*(1), 33–42. doi:10.1016/j.fcr.2007.01.003

Yang, C., & Everitt, J. H. (2012). Using spectral distance, spectral angle and plant abundance derived from hyperspectral imagery to characterize crop yield variation. *Precision Agriculture*, *13*(1), 62–75. doi:10.1007/s11119-011-9248-z

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture*, *13*(6), 693–712. doi:10.1007/s11119-012-9274-5