APPLICATION OF SATELLITE-BASED INDICES FOR PRECISE AGRICULTURAL LAND MANAGEMENT AND ESTIMATION OF EVAPOTRANSPIRATION

by

Usman Ali

Submitted in partial fulfillment of the requirements

for the degree of Master of Science

at

Dalhousie University

Halifax, Nova Scotia

July 2022

DEDICATION

My thesis is dedicated to my loving parents, Ghulam Rasool and Khalida Parveen, who are nearest to my heart.

Author

Usman Ali

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Remote sensing has become the greatest data source for large-scale studies due to the availability of remotely sensed data from a variety of sensors on various platforms with a wide range of spatial and temporal resolutions. Landsat-8, Sentinel-2A, and Planet imaging data were used in this study to generate LULC maps, crop maps, and crop evapotranspiration (ET_c) maps using machine learning methods. The outcomes of the first objective proved that using Landsat-8 and Sentinel-2A-based indices can reduce the need for ground truth data for LULC mapping. Without relying on satellite data, the results of this study also revealed that random forest (RF) is a superior classifier for LULC mapping than k-nearest neighbour (K-NN) and k-dimensional tree (KD-Tree). The results of the second objective showed that when multitemporal NDVI data was merged with multitemporal Planet imagery, the overall accuracy of crop maps using support vector machine (SVM) and decision tree (DT) increased. According to the findings of the third research, created ET_c maps with a resolution of 3 m can aid farmers in precisely estimating water crop requirements.

LIST OF ABBREVIATIONS

AAFC Agriculture and Agri-Food Canada

ANN Artificial Neural Network

DSS Decision Support System

DT Decision Tree

DVI Difference Vegetation Index

ET_c Crop Evapotranspiration

ET_r Reference Evapotranspiration

FAO Food and Agriculture Organization

GIS Geographical Information System

K_c Crop Coefficient

K-NN K-Nearest Neighbor

KD-Tree KD-Tree

LISS Linear Imaging Self Scanning

LULC Land Use Land Cover

MODIS Moderate Resolution Imaging Spectroradiometer

NDBI Normalized Built Up Index

NDVI Normalized Difference Vegetation Index

NDWI Normalized Difference Water Index

PEI Prince Edward Island

RF Random Forest

RTK Real Time Kinematics

SAM Spectral Angle Mapper

SAVI Soil Adjusted Vegetation Index

SNAP Sentinel Application Toolbox

SVM Support Vector Machine

U1 Urban Index

ACKNOWLEDGEMENTS

I express my thankfulness to "Almighty Allah," who is most gracious; he has blessed me with potential, interest, hard work, and the capacity to finish my studies. During my research study, I want to express my gratitude to my supervisor, **Dr. Travis Esau**, for his professional guidance and motivation. Thank you for your insightful insights, never-ending inspiration, constructive criticism, intellectual direction, and undivided attention during these studies. I gratefully acknowledge his invaluable assistance in completing my project. My heartfelt gratitude to **Dr. Aitazaz Farooque** and **Dr. Qamar Zaman**, members of my committee. I appreciate their assistance and encouragement on this journey.

I'd like to thank **Dr. Aitazaz Farooque** for his help and supervision throughout the project. I'd like to thank **Dr. Qamar Zaman** for his constant support throughout the project. I'd like to express my deepest gratitude to **Dr. Farhat Abbas** for his assistance in raising the quality of this project. Hanmbal Khan, Saad Aleem, Matthieu, and the overall Mechanized System research team deserve special recognition for their help and assistance.

I want to convey my profound gratitude to my parents**, Mr. and Mrs. Ghulam Rasool**, for their moral support in ensuring my success and for sacrificing their comforts for me to acquire this degree of study. I'd like to use this occasion to express my respect and gratitude to my brothers and sister for their unwavering support, love, and constant prayers for my success.

Usman Ali

CHAPTER 1: INTRODUCTION

1.1 Background

A decision support system (DSS) is needed to manage natural resources for sustainable and advanced agriculture under the concept of precision agriculture. Such DSS may include the use of land use and land cover (LULC) maps that are not only needed for natural resource management but also for crop yield estimation (Doraiswamy et al., 2004; Cihlar, 2000). LULC maps help to identify effective strategies to deal with the challenges of economic policies and environmental studies (Topaloğlu et al., 2016). In previous attempts, agricultural land use information was mostly updated by the farmers and land surveys to identify the crop type and land cover change. Such conventional methods provide detailed information about soil and its cover but are expensive due to labour costs for collecting ground truth data (Manuel, 2007). A low-cost LULC mapping technique is always desirable, especially for large areas. The use of multiple vegetation indices has the potential to develop an economic mapping methodology. This may be accomplished by creating the training vectors for LULC mapping (Muavhi, 2020). This approach can limit the need for ground truth data for LULC mapping.

Preparing the crop maps is challenging as the discrimination between different crops is problematic due to several factors, such as the same development patterns, similarity between other crops, and variability between the same crops and cloudy conditions (Peña-Barragán et al., 2011). Numerous vegetation indices are used to increase the mapping accuracy, but the normalized difference vegetation index (NDVI) and normalized difference water index (NDWI) are commonly used in crop mapping (Hao et al., 2015; Peña et al., 2014).

Field crops including, barley, oats, wheat, canola, corn, soybean, and potato are grown in Prince Edward Island (PEI), Canada. Agriculture and Agri-Food Canada (AAFC) used Landsat-8 and Sentinel satellite data to prepare crop inventory maps at 30 m resolution in conjunction with machine learning algorithms. In 2020, the overall accuracy of 85 % was achieved by AAFC in PEI. Researcher's documents that the accuracy of crop maps can be increased by using the high-resolution satellite data alone and in conjunction with vegetation indices. Planet satellite takes the earth images at 3 m resolution. The developed higher accuracy maps help the government to estimate the crop acreage and distribution. The crop maps are used in various models to estimate the crop yield and help to prepare the crop evapotranspiration maps (ET_c) for quantification of the actual crop water requirement.

Potatoes and their products are the primary agricultural commodities of the Island. Potato is the world's fourth most important food crop (Zhang et al., 2017). About 34,600 ha of the total land area (566,560 ha), is reserved for potato production (Agriculture and land department PEI, 2019). The potato industry contributes $1 billion annually to the PEI economy (Mitch Macdonald, 2020). Contributing about 10.8% to the provincial gross domestic product (GDP) and 12% of the jobs for the local islanders each year (Agriculture and Agri-Food Canada, 2017). These figures can be further improved through adopting precision agriculture techniques and good agriculture management practices. Soils of PEI are sandy loam that contains moderately high silt content and are well-drained, therefore, suitable for potato production (SOMS, 2018). The nature of such soils and the future climate change prediction would require irrigation water management for potato crops in addition to other good agriculture practices. An up-to-date crop type map is always required to design the irrigation schedules of a particular crop at a regional scale (El-Magd & Tanton, 2003).

Reference evapotranspiration (ET_r) is used to calculate crop ET_c is considered as actual crop water requirement and is affected by weather and actual crop conditions (Adamala et al., 2016). Several researchers have used remote sensing techniques to estimate ET_c at a local and regional scale and estimate the crop coefficient based on vegetation indices (Lei & Yang, 2014; Kjaersgaard et al., 2011). The Penman-Monteith equation has a standard method to estimate ET_r and is considered a point-based method (Allen et al., 2005). The Hargraves method is an alternative to estimate the ET_r when climatic parameters are limited.

1.2 Knowledge Gap

The supervised machine learning algorithms achieve higher classification accuracy compared to the unsupervised machine learning algorithms. Desai and Umrikar (2012) stated that different machine learning classification algorithms achieve different accuracies based on different types of satellite images. For LULC maps, ground truth data are required for training the machine learning algorithms and validating the results. The collection of ground truth data is time-consuming and expensive to train the machine learning algorithms for classification. In this research, the proposed framework, the remote sensing indices, will be used to create the training dataset for main classes, including agriculture, forest, bare soil, and built-up area.

The crop type mapping with optical satellite imagery is affected by weather conditions (Jiang et al., 2019). The integration of optical multitemporal vegetation indices such as NDVI, NDWI with high-resolution satellite imagery can be used to increase crop mapping accuracy (Abubakar et al., 2020; Forkuor et al., 2014). AFFC prepared the crop maps for Canada, at 30 m resolution. This study will explore the potential of Planet imagery for potato crop mapping in PEI using machine learning algorithms.

The calculation of ET_c is critical for optimal water management (Adamala et al., 2016). The point-based approaches to computed ET_c are time-consuming and costly (Allen et al., 2005). Several studies have employed remote sensing techniques to determine ET_c on a local and regional scale, as well as to calculate the crop coefficient using vegetation indicators (Lei & Yang, 2014; Kjaersgaard et al., 2011). Using remote sensing methods, Reyes-González et al. (2018) and Adamala et al. (2016) computed ET_c of wheat and corn for West Bengal and Northern Mexico. An exhaustive literature search revealed no in-depth study using precision agriculture research tools such as remote sensing and artificial intelligence to predict ET_c for potato crops on PEI.

1.3 Objectives

The specific objectives of this research study are:

1. Limiting the collection of ground truth data for land use land cover maps with machine learning algorithm
2. Integration of multitemporal Planet imagery and NDVI for improving the crop mapping in Prince Edward Island, Canada: A case study
3. Satellite based crop evapotranspiration mapping to access the variation concerning the crop development

CHAPTER 2: LIMITING THE COLLECTION OF GROUND TRUTH DATA FOR LAND USE LAND COVER MAPS WITH MACHINE LEARNING ALGORITHMS

This chapter was published in the International Journal of Geo-Information, Volume 11, Issue 6.

ABSTRACT

Land use and land cover (LULC) classification maps help understand the state and trends of agricultural production and provide insights for applications in environmental monitoring. One of the major downfalls of the LULC technique is inherently linked to its need for ground truth data to cross-validate maps. This paper aims at evaluating the efficiency of machine learning in limiting the use of ground truth data for LULC maps. This was accomplished by (1) extracting reliable LULC information from Sentinel-2 and Landsat-8 images, (2) generating remote sensing indices used to train machine learning algorithms and (3) comparing the results with ground truth data. The remote sensing indices that were tested include the difference vegetation index (DVI), the normalized difference vegetation index (NDVI), the normalized built-up index (NDBI), the urban index (UI), and the normalized bare land index (NBLI). Extracted vegetation indices were evaluated on three ML algorithms, namely, Random Forest (RF), K-Nearest Neighbour (K-NN), and K-Dimensional Tree (KD-Tree). The accuracy of these algorithms was assessed with standard statistical measures and ground truth data randomly collected in Prince Edward Island, Canada. Results showed that high kappa coefficient values were achieved by K-NN (82% and 74%), KD-Tree (80% and 78%), and RF (83% and 73%) for Sentinel-2A and Landsat-8 imagery, respectively. RF was a better classifier than K-NN and KD-Tree and had the highest overall accuracy with Sentinel-2A satellite images (92%). This approach provides the basis for limiting the collection of

ground truth data and thus reduces the labour cost, time, and resources needed to collect ground truth data for LULC maps.

Keywords: Remote sensing indices, machine learning, ground truth data, LULC mapping, Satellite imagery

2.1 Introduction

Land use and land cover classification is the most widely researched topic in the remote sensing field as it provides valuable information for urban planning, resource management, environmental monitoring, and agricultural mapping (Nguyen et al., 2018). LULC classification can be used to highlight historical trends or provide evidence-based tools in decision-making for resource management (Burkhard et al., 2012). For several years, satellite imagery has been used in LULC classification in a variety of statistical and empirical methods. Unfortunately, these methods have several limitations on accuracy assessment, as each satellite has different spectral, temporal, and radiometric resolutions (Phiri & Morgenroth, 2017). Recently, the data science and remote sensing communities have successfully achieved higher accuracy due to the launch of new satellite constellations and machine learning algorithms (Abdi, 2020). Furthermore, free access to data from earth observation satellites, including Sentinel-2 and Landsat-8, has created competition among big data scientists to increase classification accuracy with new classification algorithm approaches.

Population growth is a primary driving factor for LULC transformation (Showqi et al., 2014). As a region's population grows, so does the demand increased for built-up space and food, while other land cover types such as barren land, agricultural, and forest decline. Food consumption is increasing due to the world's population more than doubled between 1961 and 2016 (United Nations Department of Economic and Social Affairs_Population Division, 2021).

Cropland covers about a third of the land while grazing pastures and meadows cover the remaining two-thirds. Farmland area per capita has decreased substantially from 0.45 hectares per capita in 1961 to 0.21 hectares per capita in 2016. Urban areas expand outwards when agricultural and other natural land cover types are transformed into developed areas. Rapid urbanization triggered rapid economic growth and land-use changes, raising the demand for efficient natural resource management (P. Li & Moon, 2004). Increased population induced rapid urbanization, and the resulting economic growth and land-use changes heightened the demand for efficient natural resource management (Vajda & Santosh, 2017). If the land cover changes are not continuously monitored with an increasing population, it may negatively impact the environment and resource management of the area (Mohan et al., 2011). An up-to-date and cost-effective method is required to prepare the LULC map for its resource management and planning and is helpful to meet the increased food demand.

National and regional LULC maps like the AAFC Annual Crop Inventory are produced yearly and help to understand the state and trends in agriculture production in Canada. The method commonly used for preparing these maps consists of collecting ground truth data and training a classifier (Serra et al., 2009). Although this technique can be very accurate, the process of collecting reference data requires extensive planning, time, and important financial resources (Li, et al., 2017). In addition, if ground truth data cannot be collected for a year due to unforeseeable

events, the creation of LULC maps can be compromised and create a gap in the data. This situation occurred in 2020 when the AAFC Annual Crop Inventory could not be completed in Nova Scotia due to COVID-19 travel restrictions. These restrictions prevented ground data collection, making it impossible to define some agricultural classes (Fisette et al., 2014). Although ground truth data remain essential to validate LULC maps, this situation highlights the importance of developing new methods that would limit the number of ground truth data for LULC maps generation.

The quantity of training and validation data is another factor that can influence the classification accuracy of machine learning algorithms (Foody et al., 2006). For example, machine learning algorithms perform better with larger training data sets than with smaller data sets (Lu & Weng, 2007). The minimal number of training samples should be 10 times the number of variables, according to a long-standing 'rule of thumb' in machine learning (Piper, 1992). Unfortunately, there appears to be no comparable guidance in the literature regarding the minimum number of samples required for machine learning classification. According to the (Heydari & Mountrakis, 2018) classification method, the amount of input variables, and the size and spatial heterogeneity of the mapped area can all influence the number of training samples required in classification. Large, precise training data sets are required, according to the general conclusion (Huang et al., 2002). However, in the realm of applied remote sensing, training and validation data were both costly and scarce (Ramezan et al., 2021). As a result, most remote sensing studies used a single

fixed data set to test the accuracy of machine learning algorithms (Maxwell et al., 2018; Jamali, 2019).

Several supervised and unsupervised algorithms have been developed to process satellite images for LULC classification. Unsupervised classification algorithms use automated statistical algorithms to separate LULC classes without training data. Vishwanath et al. (Vishwanath et al., 2016) used an unsupervised algorithm to distinguish the land cover features in remotely sensed images. They used the K-means cluster algorithm to classify the LULC classes effectively. In another study, Shivakumar and Rajashekararadhya (2018) tested the maximum likelihood classifier for mapping the LULC classes under different scenarios.

Supervised classification algorithms are very efficient at mapping LULC features. The supervised classification method involves the representative samples from each predefined class, followed by the training of algorithms to learn about LULC classes for efficient classification. Related work by Mather and Tso and Abbas and Jaber (Mather & Tso, 2016; Abbas & Jaber, 2020) documented that supervised classification algorithms are likely to perform better than unsupervised classification algorithms.

LULC maps require an appropriate classification algorithm to solve real-world problems with high accuracy (Desai & Umrikar, 2012). Several studies have shown the potential of machine learning and statistical algorithms in LULC classification. For example, Desai and Umrikar (2012)

tested two supervised classifiers, namely maximum likelihood and minimum distance for LULC classification using Landsat imagery. The maximum likelihood classification of Landsat data had higher accuracy than the minimum distance method. Nguyen et al. (2021) used the ground truth data to train the machine learning algorithms for LULC classification. Other studies by Jia et al. (2014) tested support vector machine (SVM) and maximum likelihood algorithms on Landsat-8 imagery. The maximum likelihood classifier results were more accurate than the results of SVM. Over time, more advanced algorithms have been used in LULC classification, including decision trees (DT) and random forest (RF). Thanh Noi and Kappas (2017) tested RF, k-nearest neighbour (K-NN), and SVM using training sample sizes generated from Sentinel imagery. Results from the study showed that SVM had the highest accuracy and the least sensitivity to the size of the training sample. However, K-NN and RF classifiers attained higher accuracy with a large training size compared to SVM.

The literature also reveals that every classification method performs differently depending on the types of satellites used to capture images. For example, Jia et al. (2014) compared Landsat-7 and Landsat-8 using similar algorithms and found that the latter showed higher accuracy than the former satellite. Similarly, Ali et al. (2018) recorded higher accuracy on ALOS-2 dual-polarization bands than the Landsat-8 optical imagery data with a maximum likelihood classifier. Clerici et al. (2017) tested Sentinel-1 and Sentinel-2 satellite imagery to enhance mapping

accuracy and found higher accuracy for Sentinel-2 data than Sentinel-1 data in conjunction with the SVM algorithm. Above mentioned results proved that the accuracy of LULC maps depends on the choice of satellite and the classification algorithms used.

Due to the high cost associated with the collection of ground truth points and the heightened demand for efficient natural resource management, the objective of this study is to evaluate the efficiency of machine learning algorithms in limiting the use of ground truth data for LULC maps. This will be accomplished by extracting LULC information from Sentinel-2 and Landsat-8 satellite images and by generating remote sensing indices used to train machine learning algorithms. The results of this paper are divided into three parts. First, results from the machine learning algorithms were evaluated against ground truth data. Second, standard statistical measures were used to evaluate the performance of each machine algorithm. Third, algorithms were compared to each other to understand their performance better.

The paper has been divided into two main sections i.e., the material and method and the results and discussion section. The material utilized in this investigation and its processing details were mentioned in the material and method section. The accuracy of the machine learning algorithm is examined and compared in the result and discussion section with the findings of prior studies.

2.2 Materials and Methods

2.2.1 The Study Area

The study area consists of PEI one of Canada's smallest provinces with a land area of approximately 5,669 square kilometres (Figure 2-1). In 2019, the province had a total population of 157,262, which represents less than 0.5% of Canada's total population (Government of Prince Edward Island, 2020). The climate on the Island is mild and strongly influenced by the warm waters of the Gulf of St. Lawrence (Department of Environment, n.d.). PEI has a wide variety of landscape use, including forests, agriculture, meadows, water, wetlands, and urban areas.

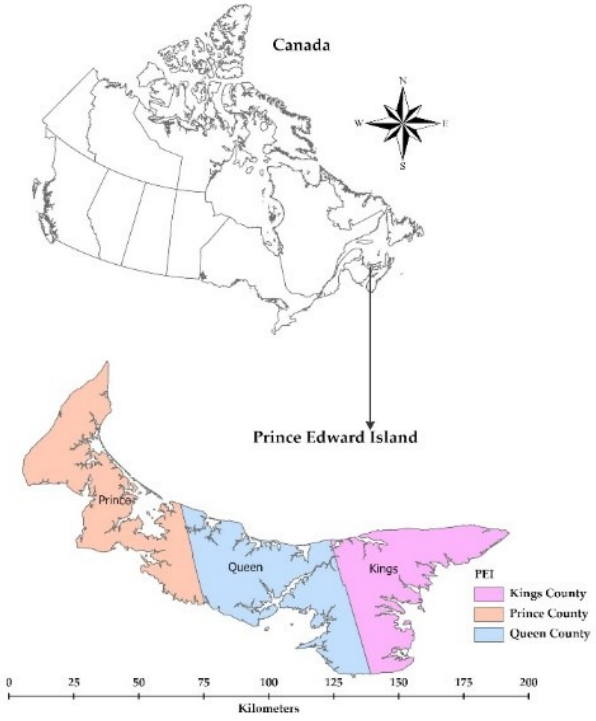


Figure 2-1. Prince Edward Island, Canada

2.2.2 Data Acquisition

Two types of satellite images were evaluated since literature reveals that classification methods perform differently on different types of satellite imagery (Asokan et al., 2020). A total of seven satellite scenes were acquired from the USGS website from July 7th to July 28th, 2019 (Table 2-1).

The Sentinels satellites are a constellation that consists of two twin satellites, Sentinel-2A and Sentinel-2B. When these satellites operate simultaneously from the same orbit, phased at 180° to each other, they can monitor the variability in land surface conditions every 5 days (Drusch et al., 2012). Sentinel-2 satellites acquire optical imagery at a resolution ranging from 10 to 60 meters depending on the spectral bands. The satellite coverage limits are between 56^{o} latitude South and 84^{o} longitude North with a swath width of 290 km.

Table 2-1. List of satellite images used in the study.

Satellites	Number of bands	Resolution (m)	Acquisition date (dd/mm/yyyy)	Path-row/tile (Number)	Cloud Cover (%)
Sentinel-2A	13	10-60	26-07-2019	T20TLS	≤10
			16-07-2019	T20TMT	≤10
			20-07-2019	T20TNS	≤10
			28-07-2019	T20TMS	≤10
Landsat-8	11	15-100	26-07-2019	008-028	≤10
			26-07-2019	007-028	≤10
			07-07-2019	007-027	≤10

The Landsat-8 satellite is also an Earth observation satellite equipped with two payloads that collect 11 spectral bands with a spatial resolution ranging from 30 to 100 m. Landsat-8 was selected due to its enhanced thematic mapper in the range of visible bands compared to other

Landsat satellites (Jia et al., 2014). The Landsat-8 has improved capabilities from the previous generation due to the addition of new spectral bands in the blue spectrum, the use of two new thermal bands and an enhanced duty cycle that has increased the daily image collection capacity of the satellite (Richardson & Wiegand, 1977).

The Landsat 8 satellite scenes were selected with the lowest cloud cover available to reduce scattering and absorption of light in the atmosphere (Table 2-1). Additionally, the satellite scenes were taken from the collection-1 level-1, which was already geometrically and radiometrically corrected.

2.2.3 Data Preparation

The Sentinel-2A and Landsat-8 images were processed using the Sentinel Application Toolbox version 8.0.0 (SNAP). All Sentinel-2A and Landsat-8 satellite image bands were resampled in SNAP using the nearest neighbour method into 20 and 30 m resolutions, respectively. The resampled images were mosaicked to cover PEI's provincial boundary using SNFAP built-in raster mosaicking tool. Three Landsat 8 scenes were mosaicked to cover the entire Island. Two of these scenes were collected on 26 July 2019, and the other one was acquired on 19 July 2019. The satellite images were reprojected to a local coordinate system, imported in ArcGIS Pro, and used to create training data for the LULC maps.

2.2.4 Remote Sensing Indices and LULC Classes

Vegetation indices such as NDVI or soil adjusted vegetation index (SAVI) can be obtained from remotely sensed data. Vegetation indices are simple to generate from multispectral satellite imagery and effective algorithms for evaluating vegetation cover quantitatively and qualitatively. Similarly, urban index, such as normalized built-up index (NDBI) can be used to identify urban

features on satellite images. In the hands of trained geospatial analysts, remote sensing indices can highlight different types of land cover and can be particularly useful to train classifiers used in LULC maps.

In this study, difference vegetation index (DVI) and NDVI were used to identify vegetation cover (P. Li & Moon, 2004). Since agriculture and forest have similar values in both indices, the barren lands were identified using the normalized bare land index (NBLI) index was used to overcome this issue Figures 2-2 (A1,5 & B1,5). The NBLI index is effective in highlighting soil composition and is helpful to differentiate agriculture from forested areas Figure 2-2(A5 & B5) (Li, Wang, Zhong, Zhang, et al., 2017). Urban features were identified with the NDBI and built-up index (UI). These indices were used because they distinguish barren land from urban features (Li, Wang, Zhong, Zhang, et al., 2017). Results of the NDBI index presented in Figure 2-2 (A3 & B3) showed that some pixels representing urban areas on the Landsat-8 and Sentinel images were mixed with bare land features. This issue was resolved by using the UI index since urban features can be identified with more precision in Figures 2-2 (A4 & B4). The minimum and maximum values of vegetation indices between 0 and 1 were shown in the legend in Figures 2-2. To construct the training samples for LULC mapping, maximum value pixels were used in each index.

Table 2-2. Remote sensing indices for highlighting LULC classes

Type	Index	Formulas	References
Vegetation Index	DVI	$NIR - Red$	(Richardson & Everitt, 1992)
	NDVI	$\frac{NIR - Red}{NIR + Red}$	(Tucker, 1979)

Urban index	NDBI	$\frac{SWIR - NIR}{SWIR + NIR}$	(Zha et al., 2003)
	UI	$\frac{SWIR2 - NIR}{SWIR2 + NIR}$	(H. Li, Wang, Zhong, Su, et al., 2017)
Barren land index	NBLI	$\frac{Red - TIR}{Red + TIR}$	(H. Li, Wang, Zhong, Su, et al., 2017)

Four LULC classes, namely agriculture, urban, barren land, and forest, were identified in the study area. were used to delineate the LULC classes. These indices were used for extracting the training samples for LULC classification in ArcGIS Pro. A total of 2000 training samples, 500 samples for each class, were created to train the classifier. The sample size was determined to be large enough since it adequately cover the entire study area without exhausting the classifier computing power.

Table 2-3. Description of land use land cover classes used in this study.

LULC Class	Description
Agriculture	Cultivated land, crop fields, vegetable fields
Urban	Residential, commercial, industrial, mix urban, other urban
Barren Land	Exposed soil, construction site, fallow land
Forest	Deciduous forest and mix forest, shrubs, and other

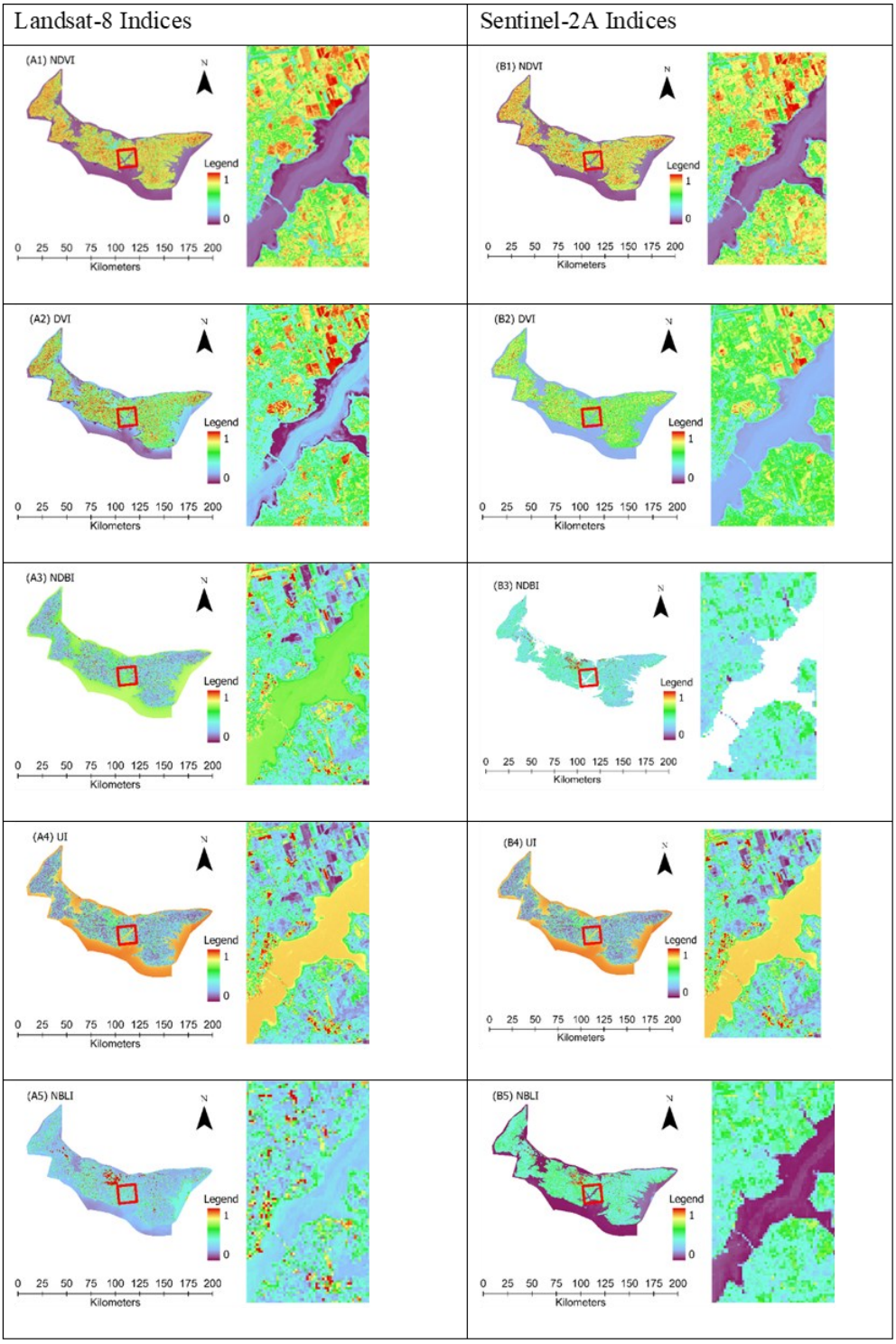


Figure 2-2. Sentinel-2A and Landsat-8 Indices are used to classify the agriculture, urban, forest, and barren land cover. The right portion of each panel shows the details of the selected portion.

2.2.5 Machine Learning Algorithm

2.2.5.1 Random Forest Classifier

The RF classifier is a combination of tree predictors with each tree depending on an independently sampled random vector value with similar distribution in all trees (Figure 2-3) (Breiman, 2001). Boosting and bagging are two ensemble methods capable of squeezing additional predictive accuracy out of classification algorithms. Bagging algorithms are used to reduce the complexity of the models that overfit the training data; while the boosting algorithm increases models' complexity. The training samples, which are not used in the training sample were included in the evaluation and were referred to as 'out of bag' samples (Abdi, 2020). In addition, the RF classifier is easy to use since it only uses two parameters (e.g., number of variables at each node and number of trees), which is not sensitive to the parameter value (Liaw & Wiener, 2002). The number of trees and predictors in RF classification are vital parameters to achieve the highest accuracy possible. For assessing the accuracy of the current RF output, these parameters were set at 50 for the number of trees and a maximum number of tree depth and sample per class were set as 30 and 500, respectively.

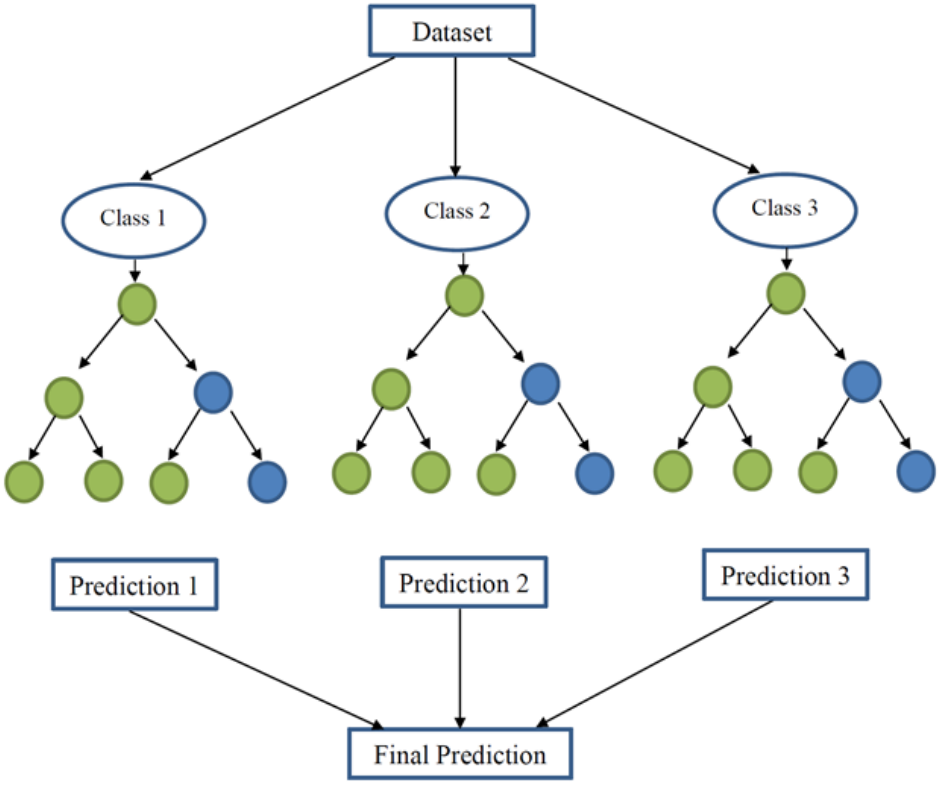


Figure 2-3. A pictorial view of the Random Forest working principle.

2.2.5.2 K-Nearest Neighbour

The K-NN is a supervised machine learning algorithm that can be used to solve classification and regression problems. It was first discussed in an unpublished report by (Fix, 1951), followed by more detailed K-NN rules published by (Cover & Hart, 1967). It categorizes the objects based on the nearest neighbour class. The major deciding factor in the classification task is the number of neighbours (k) used to classify an object (Figure 2-4). Small k values indicate relatively inaccurate results while higher k values indicate a more credible result (Karegowda et al., 2012). Through trial and error, the optimal k value was found and set to k = 20.

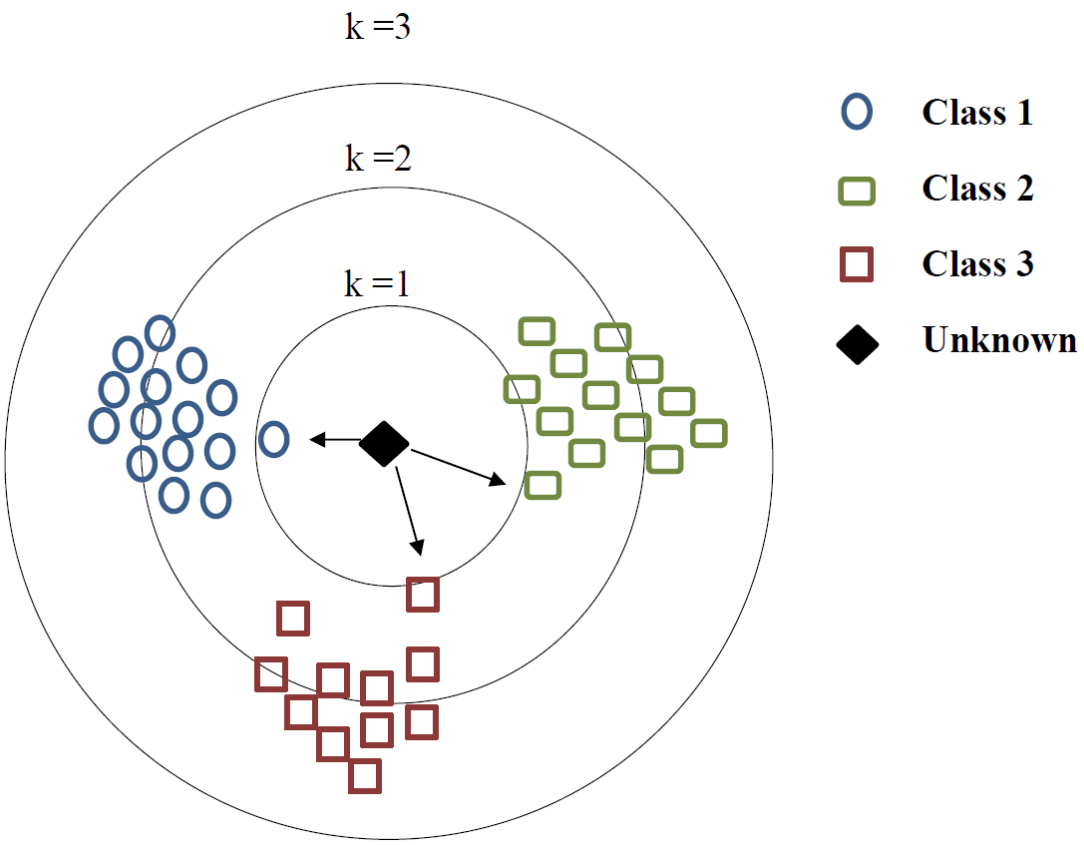


Figure 2-4. A pictorial view of the K-Nearest neighbor working principle.

2.2.5.3 K Dimension Tree

KD-Tree is the most common binary algorithm used for the nearest neighbour algorithm family. In KD-Tree classifiers, the clusters are developed based on the median of x and y axes (Figure 2-5). KD-Tree categorizes points based on the projections in lower dimensions (Narasimhulu et al., 2021). For lower-dimensional datasets, the KD-Tree is designed to perform better compared to other algorithms such as Ball and Tree (Dolatshah et al., 2015). For accuracy comparison of KD-Tree on both satellites, the number of training samples was set at 2000 and the number of neighbours was set at 20. Similar to the k value in the K-NN algorithm, the optimal number of neighbours was determined through trial and error.

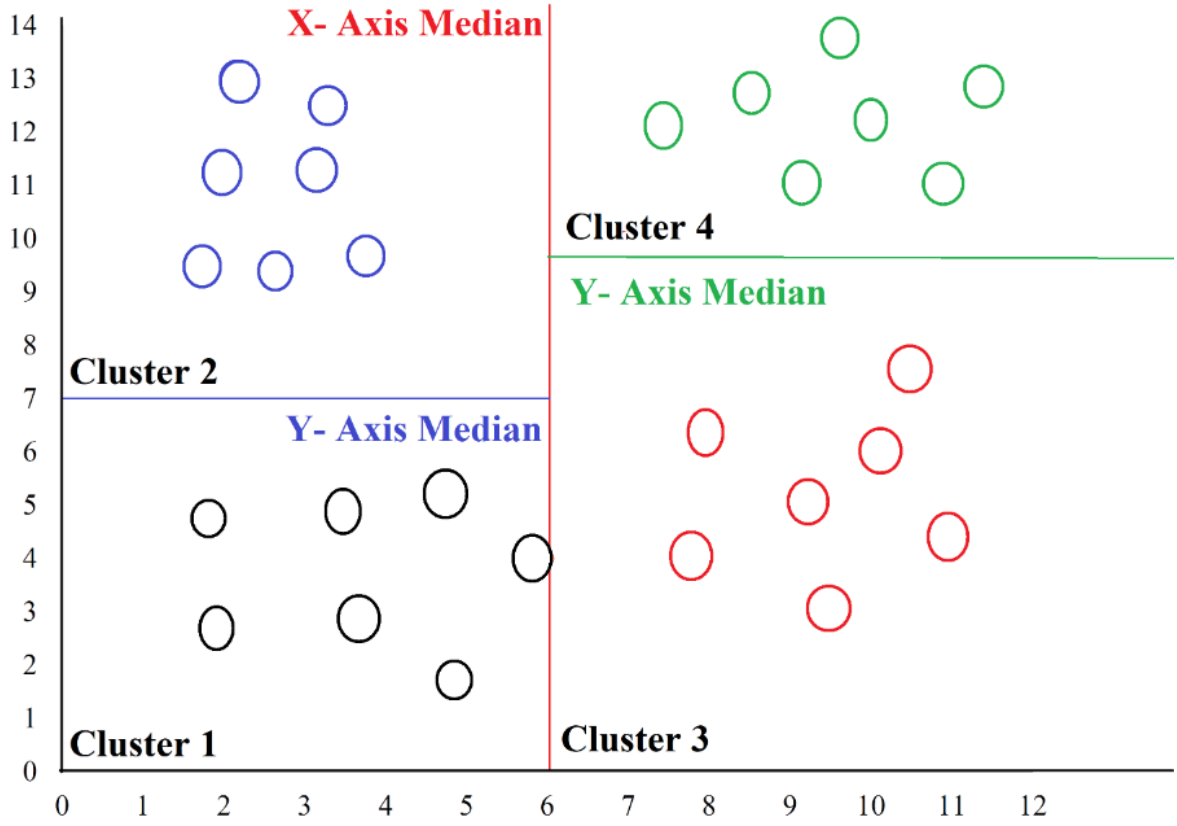


Figure 2-5. A pictorial view of the K Dimension Tree working principle.

2.2.6 Ground Truth Data for Validation and Model Evaluation Criteria

Five sites on PEI were selected to collect ground truth data. Using a Real-Time Kinematic (RTK) GPS with sub-meter accuracy, a total of 200 validation points were collected at each site. These points were equally distributed in each class, meaning that 50 points were collected per class at each site. The same ground truth data were used to validate LULC maps generated with Sentinel-2A and Landsat-8 imagery.

Several statistical indicators were used to assess the accuracy of the models. The overall accuracy of the models was used to describe the correct proportion of mapped pixels. The overall

accuracy considers that 100% of all the classified reference sites are mapped accurately (Story & Congalton, 1986). The overall accuracy was calculated using the following formula:

$$\text{Overall Accuracy (\%)} = \frac{\text{Number of correctly classified pixels}}{\text{Total number of referenced site pixels}} \times 100 \qquad (2.1)$$

Similarly, each LULC class's accuracy was determined using producer/user accuracies. Producer/user accuracy determines the real feature on the ground surface correctly shown on the classified map (Story & Congalton, 1986). According to the literature discussed in the introduction, the optimal values for the producer, user, and overall accuracy are around 80 to 85%. The producer and user accuracy were calculated using the following formula:

$$\text{Producer/User Accuracy (\%)} = \frac{\text{Correctly classified pixels in one category}}{\text{Total classified pixels in all categories}} \times 100 \qquad (2.2)$$

The Kappa coefficient is another statistical indicator to evaluate classification accuracy. Kappa evaluates how well the classification has performed compared to the randomly assigned value. Its values range from -1 to 1, with the lowest value indicating that the classification is not better than a random classification. While a value close to a positive one indicates that the classification is significantly better than the random classification (McHugh, 2012). The Kappa coefficient was calculated using the following formula:

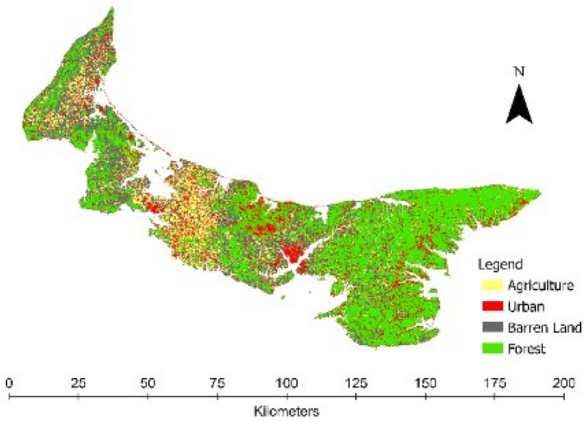
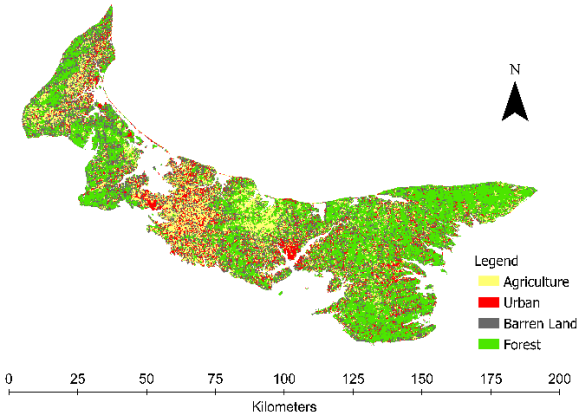
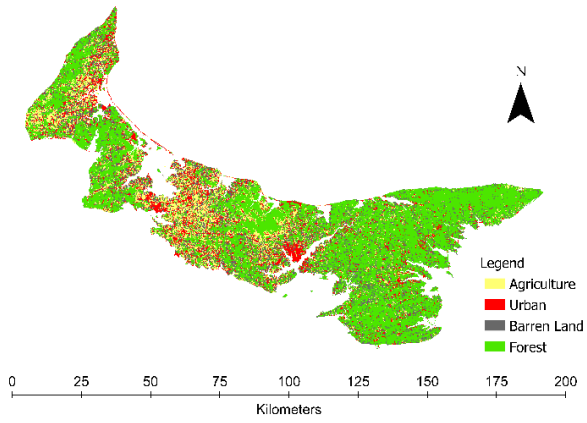
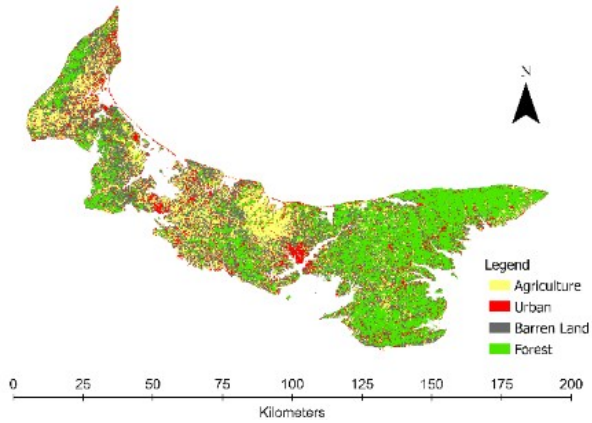
$$\text{Kappa Cofficient} = \frac{(TS - TCS) - \sum(Column\ total * Row\ Total)}{TS^2 - \sum(Column\ total * Row\ Total)} \times 100 \qquad (2.3)$$

where *TS* is the total number of samples, *TCS* is the total number of classified samples and column sum and row sum represent the total number of classified pixels for each class in each column and row, respectively.

2.3 Results

2.3.1 Land Use Land Cover Mapping Results

In prepared LULC maps (Figure 2-6), the yellow colour represents the agricultural area, the green colour represents the forest area, battleship grey represents barren land, and red represents the urban area.

Classifier	Landsat-8	Sentinel-2A
KD-Tree		
RF		

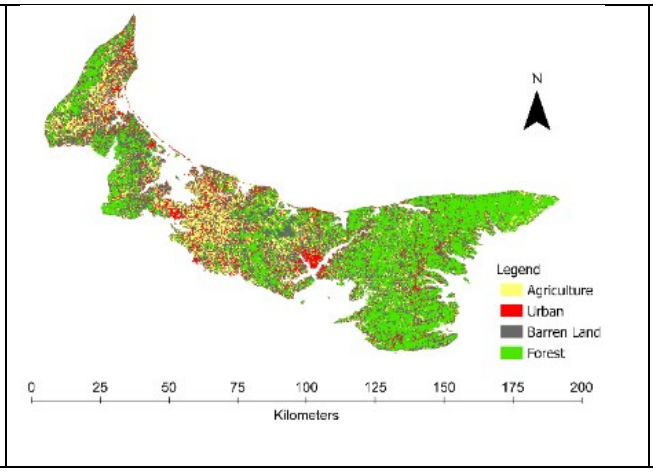
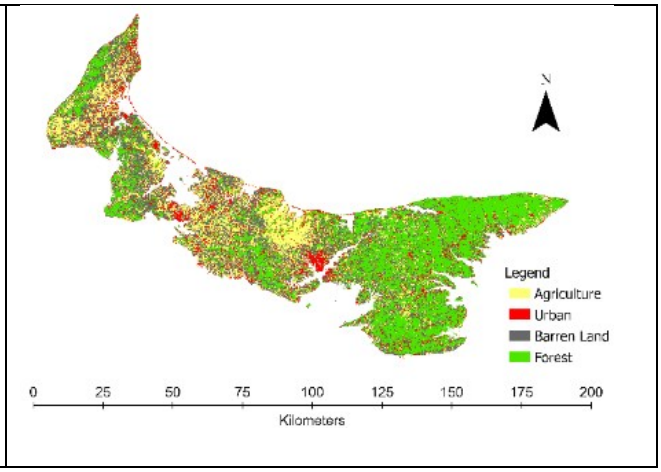
K-NN		

Figure 2-6. Remote sensing indices based LULC map prepared based on two satellite data.

From the Landsat-8 imagery, the KD-Tree classifier detected the true positives for the agriculture class, e.g., 45 out of 50 with user accuracy of 90% and producer accuracy of 79% respectively (Figure 2-7(a) and Table 4). For Sentinel-2A imagery, the highest true positives were classified by the K-NN algorithm, e.g., 47 out of 50 for the agriculture class. The RF and K-NN in Landsat-8 and KD-Tree and RF in Sentinel-2A recorded true positives for agriculture classes ranging from 38-45 out of 50 (Figure 2-7(a, d)). Interestingly, the highest and lowest true positives for the agriculture class were recorded by the K-NN algorithm. This implies that the performance of classification may be improved by using finer resolution and more refined imagery (D. Chen et al., 2004; Rao et al., 2021).

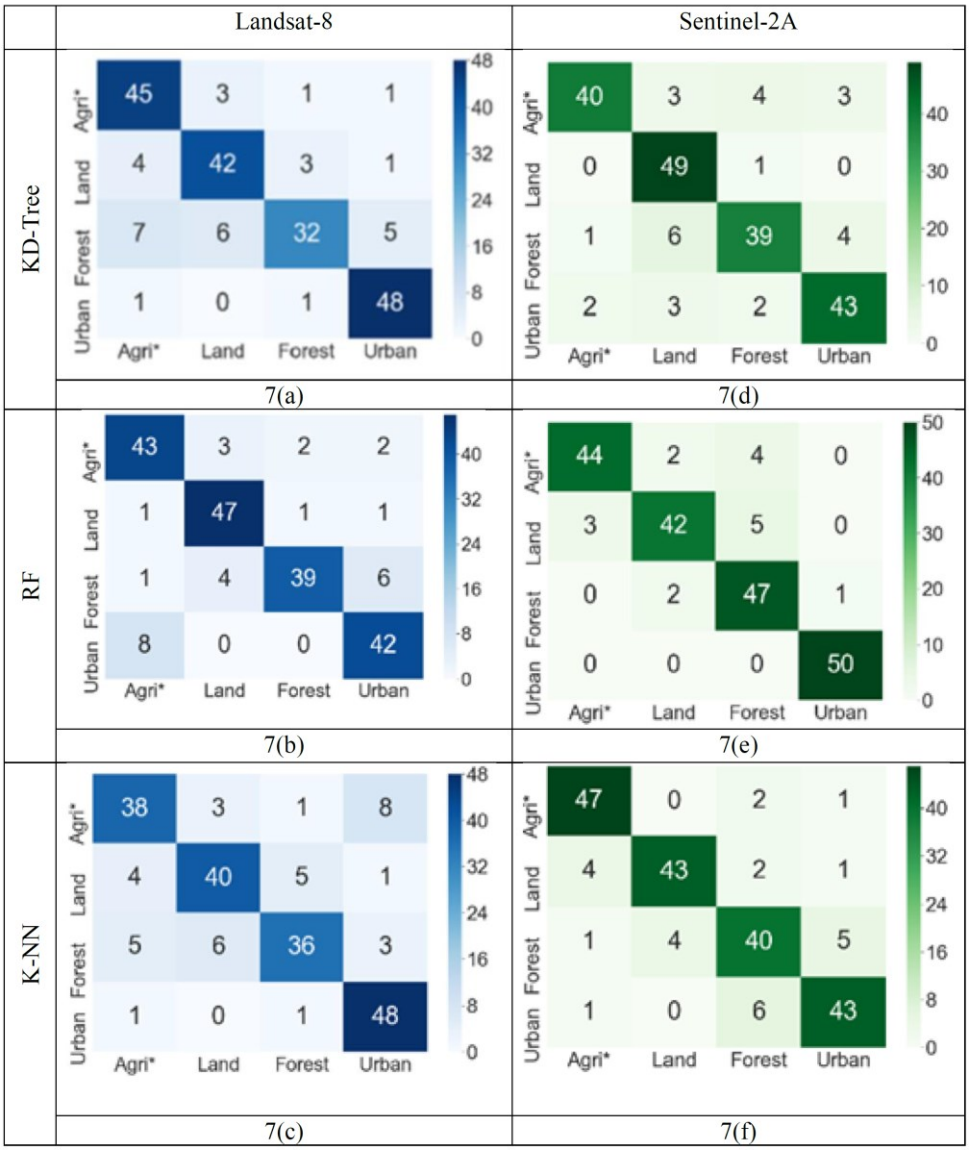


Figure 2-7. Error matrix showing correct and incorrect training vectors of the evaluation samples by each machine learning algorithm. The error matrices used to calculate the producer and user accuracy shown are in Table 2-4.

For the barren land class, the highest true positives were recorded by the KD-Tree classifier with the Sentinel-2A imagery, i.e., 49 out of 50 with a user accuracy of 98% and a producer accuracy of 80% (Figure 2-7(d) and Table 2-4) (Nguyen et al., 2020). However, the performance

of the KD-Tree classifier with Landsat-8 imagery for barren land class recorded relatively lower true positives, e.g., 42 out of 50. Similarly, the highest true positives for the urban forest class were recorded by random forest classifier with Sentinel-2A imagery. However, a relatively lower number of true positives was recorded for the forest class with Landsat-8 imagery, e.g., 32, 39, and 36 for KD-Tree, RF, and K-NN classifiers, respectively (Figure 2-7(a, b, c)). For the urban class, the highest average true positives were recorded for both satellite images. For the urban class, the RF algorithm with the Sentinel-2A satellite achieved the highest possible user accuracy (100%) compared to all other satellite-algorithm comparisons (Table 2-4). These results concur with the findings reported in the literature that mentions that the resolution, image characteristics, classification algorithms, and the need of the user affect the classification accuracy of LULC mapping (Chen et al., 2004; Nguyen et al., 2020).

Table 2-4. User and Producer accuracies for LCLU types used in this study.

Classifier	Classes	User Accuracy (%)	Producer Accuracy (%)	User Accuracy (%)	Producer Accuracy (%)
		Sentinel-2A		Landsat-8	
KD-Tree	Agriculture	80	93	90	79
	Barren Land	98	80	84	82
	Forest	78	85	64	86
	Urban	86	86	96	87
RF	Agriculture	88	94	86	81
	Barren Land	84	91	94	87
	Forest	94	84	78	92
	Urban	100	98	84	82
K-NN	Agriculture	94	89	76	79
	Barren Land	86	91	80	82
	Forest	80	80	72	84
	Urban	86	86	96	80

2.3.2 Satellite Accuracy Comparison

For Landsat 8 imagery, the algorithm's kappa coefficient was recorded as 78, 80, and 74% for KD-Tree, RF, and K-NN, respectively (Figure 2-8). For Sentinel-2A imagery, the same algorithms recorded considerably increased kappa coefficient values, i.e., 2.5, 10, and 10.8% for KD-Tree, RF, and K-NN algorithms, respectively. Similarly, the average kappa coefficient was 83.3% for the Sentinel-2A while the average was 77.3% for the Landsat-8.

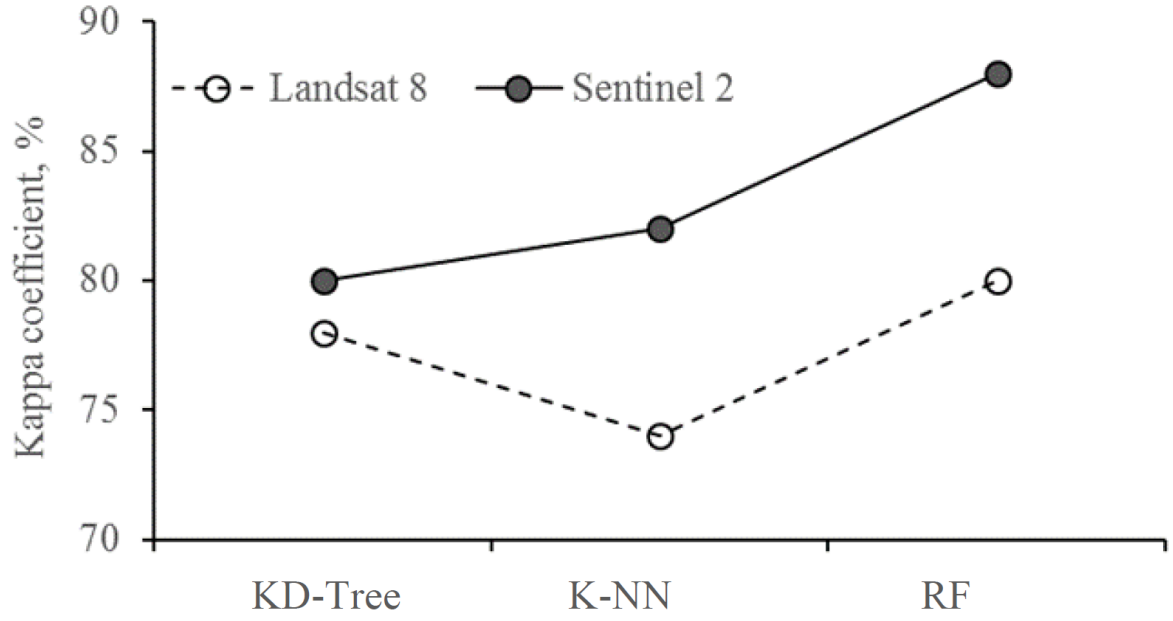


Figure 2-8. Kappa coefficient comparison of two satellites using different classifiers.

Random Forest classifier's overall accuracy was recorded as 92 and 85% for Sentinel-2A and Landsat-8 satellites, respectively (Figure 2-9). The average accuracy of the KD-Tree classifier for both satellites was recorded to be 84.5%. The K-NN achieved 86 and 81% overall accuracy for Sentinel-2A and Landsat-8 satellites, respectively. A slightly lower average overall accuracy of 83.5% was recorded for the K-NN algorithm in comparison with the KD-Tree classifier (Figure 2-9).

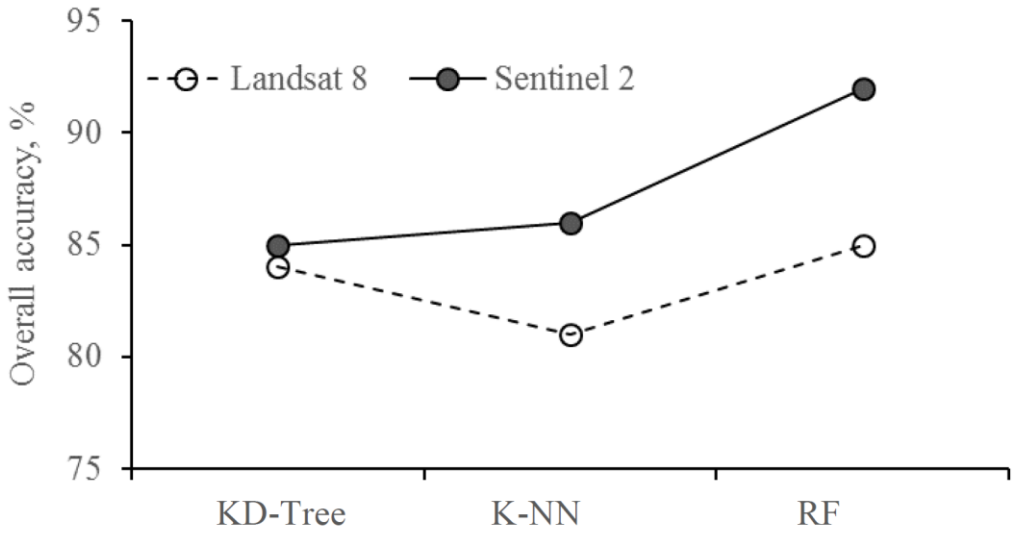


Figure 2-9. Overall accuracy comparison of two satellites using different classifiers.

2.4 Discussion

The Sentinel-2A and Landsat-8 presently operate at medium resolution at 10, 20, and 30 m. The resolutions of these two satellite bands are different. Before further processing, all Landsat-8 bands were resampled to 30 m resolution, while Sentinel-2A bands were resampled to 20 m resolution. This study presented the potential of different remote sensing indices to create the training samples for LULC mapping in PEI in conjunction with three machine learning algorithms. The Island population is increasing and major land cover classes such as forest, agriculture, barren land, and urban will be affected. These rapid changes demand more effective methods to map land cover changes and conduct resource management analyses.

The remote sensing indices, including DVI, NDVI, NDBI, UI, and NBLI, were selected to highlight agriculture, forest, barren land, and urban area. This approach to preparing LULC maps is much cheaper and faster than other classification methods traditionally used. Although, for some LULC classes, it is hard to find suitable remote sensing indices. For example, it is hard to distinguish between forests and agriculture using remote sensing indices. The NBLI was used to

overcome this problem because H. Li, Wang, Zhong, Zhang, et al. (2017) documented that the NBLI can highlight the soil composition at the pixel level, which helps distinguish between agriculture and forest (Figure 2C). The results from the experiment also verified the validity of this proposed method.

In the last step, we used the same algorithms processing conditions (same training and validation data sets) to compare the Landsat-8 and Sentinel-2A optimal data sets for LULC mapping. The comparison results indicated that the overall accuracy of each algorithm highly depends on the input data on the results. For example, the highest overall accuracy of RF 92 % showed that RF offers the best classification results for Sentinel-2A; whereas KD-Tree and K-NN overall accuracy slightly decreased. Interestingly, RF also offers the highest overall accuracy 85% for Landsat-8; likewise, KD-Tree and K-NN overall accuracy slightly decreased compared to RF. All the results mentioned above proved that the outcomes of each classifier depend on the input data set. The results proved that RF is a suitable machine learning algorithm as compared to the KD-Tree and K-NN for land cover classification without considering input data sets. Therefore, it is necessary to compare the obtained results with the literature because it offers a realistic view of this study's results.

For example, Lowe and Kulkarni (Lowe & Kulkarni, 2015) used the RF, SVM, and maximum likelihood classifier for preparing the LULC map and achieved an overall accuracy of 87, 83, and 77%, respectively. Another study from Franco-Lopez et al. (Franco-Lopez et al., 2001) prepared LULC maps with 13 classes using the K-NN algorithm and achieved an overall accuracy of 63%. These different results indicate that there are no clear rules for acceptable accuracy for any land cover type, but it depends upon the user and adoptive methodology. In any LULC classification, errors are present in the form of estimation and prediction (Salovaara et al., 2005).

So far, no clear rules have been defined on the acceptable accuracy range because different users have different concerns about classification accuracy (Olofsson et al., 2014). In addition, several factors influence the accuracy of classification such as image quality, classifier, number of classes, and number sample size (Nguyen et al., 2020). One study (Nguyen et al., 2020) used Sentinel-2 data for LULC mapping in Vietnam with RF, K-NN, and SVM algorithms, and reported the highest accuracy by RF when the Training vector size was appropriate to cover the study area. RF achieved higher accuracy than the SVM by using Sentinel-1 data of the Brazilian Amazon (Chaves et al., 2020). These results indicate that Sentinel-2A and Landsat-8 data had satisfactory performance in LULC mapping. In (Nguyen et al., 2020) recommended RF for LULC classification because of ease in parameter selection in RF. The results of this study concur with the findings of (Nguyen et al., 2020).

2.5 Conclusions

This study proposed a methodology to produce LULC maps at less cost and in a quick manner by using three machine learning algorithms (KD-Tree, RF, and K-NN) and two satellites (Landsat-8 and Sentinel-2A). Timely updated maps can help the local authorities for better resource management and land use policy decisions. The proposed methodology to develop the LULC maps with remote sensing indices can be leveraged by researchers to determine the spatial-temporal changes of LULC due to human activities. The results of this study demonstrated the potential of remote sensing indices to limit the need for ground truth data for LULC mapping. This would lower the labour cost, time, and resources required to generate LULC maps.

In this study, training samples for four classes; forest, agriculture, urban and barren land were created on behalf of indices and these training samples were used in conjunction with the machine learning algorithms for LULC mapping. Prepared LULC maps based on this proposed

methodology showed promising results when it was validated with ground truth data. The six LULC maps, produced by running the three machine learning algorithms using the same training data for the two sources of imagery, were subjected to accuracy assessment to determine the effectiveness of the algorithms. resource

Results from the study demonstrated that K-NN achieved an average kappa coefficient of 82 and 74% and high overall accuracy of 86 and 81% for Sentinel-2A and Landsat-8, respectively. In comparison, the KD-Tree had an average kappa coefficient of 80 and 78% and overall accuracy of 85 and 84% for Sentinel-2A and Landsat-8. Random Forest achieved the highest average Kappa coefficient of 83.3 and 73.3% and high overall accuracy of 92 and 85 % for Sentinel- 2A and Landsat-8 data, respectively, compared to K-NN and KD-Tree.

Further research should be conducted in two tasks; (1) the evaluation of this methodology on satellite images with higher resolution as well as refining the data by training samples for subclasses of crops such as potato, wheat, rice, maize and grasses. (2) The quantity and quality of training samples have an impact on land cover classification. By assuring quality and increasing the training sample size, classification accuracy can be enhanced. The ideal combination of training sample sizes will also be researched in the future.

CHAPTER 3: INTEGRATION OF MULTITEMPORAL PLANET IMAGERY AND NDVI FOR IMPROVING THE CROP MAPPING IN PRINCE EDWARD ISLAND, CANADA: A CASE STUDY

ABSTRACT

Crop mapping using high-resolution satellite images and machine learning techniques has increased our understanding of the agroecosystem and has been used to inform policy. High-resolution remote sensing images and vegetation indices like the normalized difference vegetation index (NDVI) are routinely used to construct precise and accurate crop maps, which are then checked using ground truth data. The application of machine learning algorithms for crop maps has recently shown promising results in boosting workflow performance, but it must be evaluated in a variety of scenarios. Using high-resolution satellite images, this study will assess the performance of SVM and DT algorithms for crop mapping on Prince Edward Island (PEI), Canada. This was achieved by acquiring multitemporal Planet satellite imagery during specific periods to cover the growth cycle of the targeted crops. SVM and DT algorithms were tested with four spectral bands and vegetation indices, namely Red, Green, Blue, NIR, and NDVI, at a spatial resolution of 3 meters. SVM and DT training and validation data were split following an 80/20 ratio, respectively. The validation data were used to generate confusion matrixes, producer accuracy (PA), user accuracy (UA), overall accuracy (OA) kappa coefficient, and F1 score for each crop type. F1 was also used to judge the accuracy of the prepared crop maps and compare the SVM and DT algorithms. Results showed no significant difference in the OA of crop maps while using RGB, and NIR Planet imagery as input to SVM and DT. The use of NDVI data combined with the four spectral bands improved the overall accuracy of crop maps using SVM and DT. When combined multitemporal Planet imagery and multitemporal NDVI data were employed, the SVM method had a 6.25 % higher OA and a 7% higher kappa coefficient than the DT algorithm.

This study shows how vegetation indicators can increase the effectiveness of the supervised machine learning method. This study's approach to precise crop mapping is appropriate for large-scale implementation across Canada and other parts of the world.

Keywords: Planet scope, Normalized difference vegetation index, Decision Tree, Support Vector Machine, High-resolution satellite, machine learning.

3.1 Introduction

Remote sensing, in combination with geographical information systems (GIS), and machine learning algorithms, is commonly used in agricultural monitoring applications such as crop mapping, crop water demands computation, etc. Remote sensing in conjunction with machine learning algorithms is considered a reliable technique to produce crop maps with higher accuracy (Government of Canada, 2015). Produced maps at a higher resolution are more beneficial to prepare the up-to-date crop inventory for better decision making such as implementing by-laws of cover rotation (Kussul et al., 2014). The accuracy of crop maps depends on numerous factors such as the resolution of satellite images, the accuracy of ground truth data, and classification algorithms used to map the crop types (Lu & Weng, 2007).

From 2009 to 2010, the Science and Technology Branch of AAFC started to prepare crop maps for Canada using Radarset-2 satellite data in combination with the machine learning algorithm DT (Fisette et al., 2014). After 2010, the Landsat-8 and Sentinel-2 satellite data were used to prepare the crop maps at 30 m resolution (Illert & Afflerbach, 2020). Using the Landsat-8 and Sentinel-2, the AAFC achieved overall accuracy between 95 to 85% for crop maps in Canada; however, PEI achieved the lowest overall accuracy of 85% (Illert &Afflerbach, 2020).

There are numerous reasons for low overall accuracy in PEI crop mapping. For example, the ground truth data used by AAFC for training and validation of crop mapping is taken from the records of crop insurance agencies. By ground survey, the missing ground-truth data that are not included in the crop insurance data are collected (Fisette et al., 2014). Due to this, Fisette et al. (2019) documented that acquiring the ground truth data from crop insurance companies is erroneous and leads to misclassification. Therefore, consistent ground truth data and higher

resolution data are currently lacking for crop mapping in Canada (Fisette et al., 2019). This study will solve the above-discussed issue using high-resolution satellite data, consistent ground-truthing, and the use of new machine learning algorithms.

Crop maps were previously developed using machine learning methods and freely accessible low to medium resolution satellite data. For example, Shelestov et al. (2017) evaluated the Landsat-8 and Sentinel-2 data in Google Earth Engine by using the artificial neural network (ANN), SVM, and RF algorithms for crop mapping accuracy and reported ANN to achieve higher accuracy. Kumar et al. (2015) tested the crop mapping accuracies with SVM, ANN, and spectral angle mapper (SAM) by using the linear imaging self-scanning (LISS IV) sensor data. They found that SVM and ANN showed higher accuracy for crop mapping than SAM. Vibhute and Gawali (2013) evaluated five hundred fields for crop mapping and concluded that the SVM classifier had higher accuracy than the RF and ANN. Kumar et al. (2017) tested the LISS IV sensor data to map wheat, lentils, mustard, pea, linseed, corn, sugar cane, and other crops with SVM and maximum likelihood classifier that SVM achieved results at higher accuracy than maximum likelihood classification algorithms. Several researchers have used coarse to medium spatial resolution multispectral satellite data originating from Landsat-8, Sentinel, and MODIS (Moderate Resolution Imaging Spectroradiometer) for crop mapping across the globe and reported moderate to reasonably acceptable mapping accuracy ranging from 80 to 85 % (Kussul et al., 2016; Zhu et al., 2017; Yi et al., 2020).

High-resolution satellite data with machine learning algorithms have recently been used for crop mapping (Yang et al., 2019). The use of higher-resolution satellite images such as SPOT (French: Satellite Pour l'Observation de la Terre; English: Satellite for observation of Earth) and

IKONOS (Greek word for ‘imaging’) results in higher mapping accuracy compared to the ordinarily available satellites such as Landsat-8 and Sentinel (Boyle et al., 2014). Nowadays, higher temporal and spatial resolution imagery is easily accessible due to the new Planet satellite (Cheng et al., 2020). It consists of 150-200 nanosatellites in the orbital, making it capable of acquiring 3 to 5 m resolution multispectral imagery daily (Houborg & McCabe, 2016). Planet imagery has already been used to detect crop sowing, field boundaries, agronomic parameters, and crop mapping in China's Jiangsu province (Li et al., 2019; Cheng et al., 2020). The novelty of the current study is that it uses consistent ground truth data, and Planet satellite (3 m resolution) imagery in conjunction with machine learning algorithms (SVM and DT) to map crops in PEI, Canada.

The vegetation indices sharpen the vegetation signals for detecting and classifying crop phenology (Xiao et al., 2002). Multitemporal vegetation indices can discriminate between crop types from the phenological spectral signatures of crops (Hatfield & Prueger, 2010). The NDVI has been widely used to enhance the classification accuracy of crop maps (Fitting & Models, 2019; Sitokonstantinou et al., 2018).

Agriculture-rich PEI lacks crop maps with improved accuracy and resolution. For improved agricultural monitoring and policymaking, the province requires high-resolution crop maps. In light of this, the first part of this research was to create high-resolution crop maps using a combination of multitemporal Planet imagery and NDVI images obtained from the Planet satellite. The performance of SVM and DT for crop mapping was assessed in the second part.

3.2 Material and Methods

3.2.1 Study Area

Prince Edward Island is a Canadian province dominated by agriculture (Figure 3-1). The island is split into three counties (i.e., Prince, Queens, and Kings) and is located between 46.5107° N and 63.4168° W. The climate of this agriculturally rich region is moderate and heavily affected by the warm waters of the Gulf of St. Lawrence, according to the (Department of Agriculture and Land, 2020). The Island receives 890 mm of rain and 290 cm of snow per year, with an average temperature of -7°C in January and 19°C in July. Its agricultural output scale is immense; for example, agriculture is permitted on 240515 hectares of a total area of 566560 ha. Forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry are the main crops on this island. Potatoes have 85,500 acres of dedicated land, whereas corn, wheat, barley, and soybeans have 17,6000 acres (Department of Agriculture and Land, 2020). To minimize overplanting of row crops and maintain soil health, the agricultural crop rotation legislation mandates a three-year crop rotation. Crop maps with a greater resolution will be useful in enforcing the crop rotation legislation.

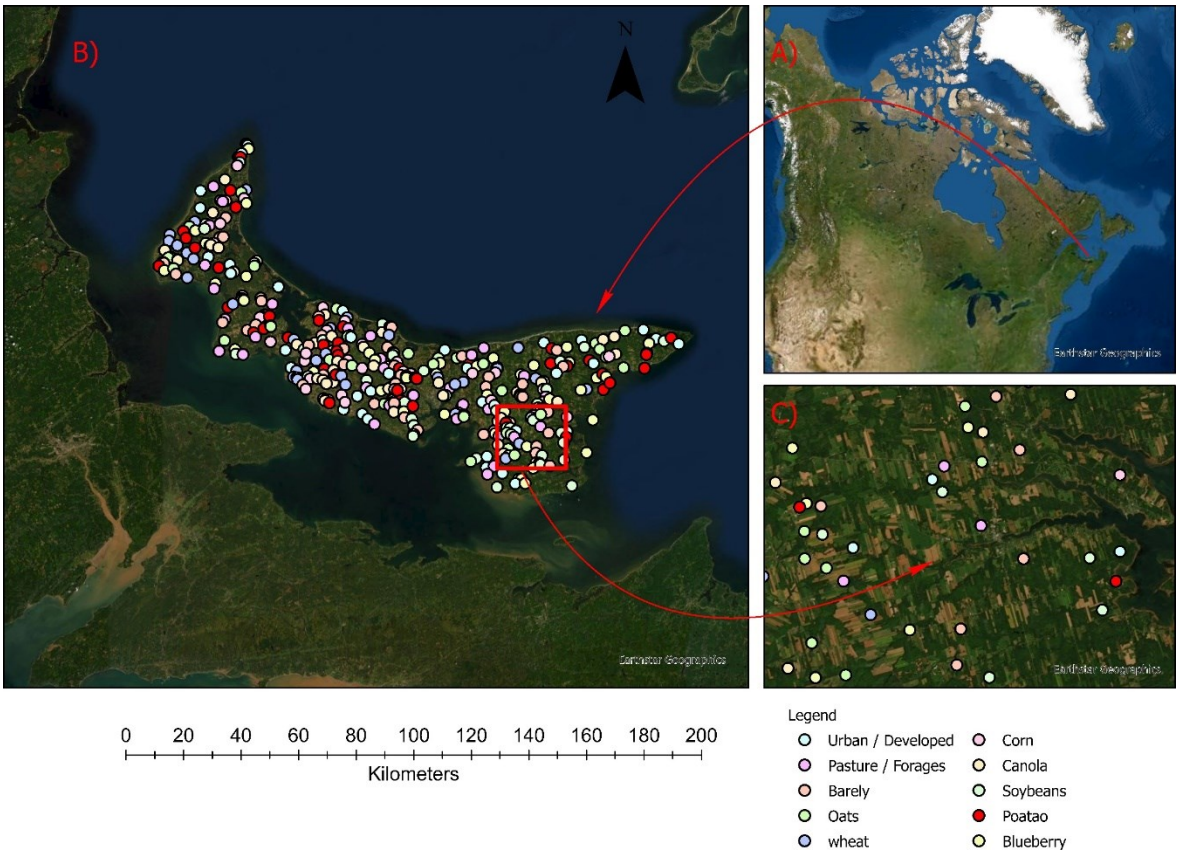


Figure 3-1. (A) Study region along with Canadian map. (B) Zoom-in image of Prince Edward Island and distribution of training and validation points overlaid on Planet satellite imagery. (C) A zoomed-in view showing the distribution of training and validation points.

3.2.2 Field Data/Ground Truth Data

During the crop growing season in the summer of 2020, a field scout was conducted to collect ground truth data for the training and validation of machine learning algorithms. Land cover and crop type classes latitude and longitude coordinates were recorded with an accuracy level of 3 m using a German terra x GPS unit (Garmin Ltd, Kansans, USA). This approach resulted in a CSV file with 2000 coordinates for urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry across PEI. These are point-based ground truth data that are not uniformly

distributed throughout the study area (Figure 3-1). The data was imported as a point shapefile into ArcGIS Pro (Environmental System Research Institute, Redland, California) and overlaid over Planet satellite imagery (Figure 3-1). The ground truth data associated with urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry are represented by the indicolite green, heliotrope, medium sand, peridot green, atlantic blue, rhodolite rose, topaz sand, lemongrass, red, and yellow yucca points in the legend of Figure 3-1. All of the data were randomly divided into two parts, with an 80% to 20% split. 80 percent of the data was utilized for classification, while 20 percent was used for validation.

3.2.3 High-Resolution Satellite Data

The Planet satellite (Planet Lab, San Francisco, California, United States) is the world's largest microsatellite constellation that captures multispectral images at 3 m spatial resolution (Cheng et al., 2020). The revisit time of this satellite is about one day to cover 340 million km^2 per day (Planet Team, 2021). Planet imagery has four spectral bands; Blue (455- 515 nm), Green (500-590 nm), Red (590-670 nm), and NIR (near-infrared; 780-860 nm). Recently, (Rao et al., 2021; Cheng et al., 2020) have used Planet imagery for crop classification and delineating crop boundaries in India and China, respectively.

To cover PEI, four monthly Planet imageries were acquired through www.planet.com from April 1, 2020, to August 1, 2020, since the majority of crops were cultivated during these months. To cover the PEI borders, almost 148 quads of Planet imageries were downloaded each time. To eliminate the scattering and absorption problems, cloud-free imageries were used. Georeferencing problems were checked in ArcGIS Pro by superimposing Planet satellite images onto satellite-based imagery to align the coordinate system with the ground coordinate system.

3.2.4 Vegetation Indices for Improving Crop Classification

To improve the classification performance of Planet images, multitemporal vegetation indices were created. Sakuma and Yamano (2020) advocated for the use of NDVI because of its significant potential for improving crop classification accuracy. One of the most important reasons for using the NDVI in this investigation is that it is well-known for highlighting crop characteristics with high classification accuracy (Sakuma & Yamano, 2020; Ji et al., 2018). Tucker (1979) proposed the NDVI formula (NIR-Red)/(NIR+Red) was used to develop multitemporal NDVI profiles for urban classes and crop types throughout the growing season. The multitemporal NDVI profile for the whole growing session was prepared using the approach described by Reyes-González et al. (2018), which involved taking the mean value of five well-represented pixels of NDVI in each class every month.

3.3 Machine Learning Algorithms

3.3.1 Support Vector Machine Algorithm

Support vector machine is a supervised machine learning algorithm used for classification and regression purposes (Mustafa et al., 2017). It helps to differentiate between classes of an object based on the hyperplane (Wang et al., 2019). In numerous studies, this algorithm has been applied for crop classification in conjunction with different satellite data (Rao et al., 2021; Mantero et al., 2004). Kumar et al. (2017) documented that SVM is important in the field of remote sensing because it can produce accurate results in the presence of limited training data. Multitemporal Planet imagery stand-alone (data set A) and in combination with NDVI (data set B) was used as input in SVM to evaluate if adding NDVI improved the classification accuracy (Table 3-1). The classification with the SVM has been performed using the ArcGIS Pro classification wizard. The

collected ground truth data were randomly split into 80% and 20% sections for the training and validation of SVM.

The SVM has only one parameter, the number of samples per class. The parameter is set through trial and error to produce better classification results using SVM (Onojeghuo et al., 2018). By adopting the trial-and-error method, samples per class (200) were selected for the optimal classification result for Planet imagery stand-alone and in combination with vegetation indices.

Table 3-1. The combinations of the multitemporal Planet imagery with NDVI data are under investigation for crop mapping classification.

Data sets	Model input	Algorithms
A	Planet imagery	SVM
B	Planet imagery + NDVI	SVM
C	Planet imagery	DT
D	Planet imagery + NDVI	DT

* Normalized difference vegetation index (NDVI), Support vector machine (SVM), Decision Tree (DT). All the data sets such as (A= Planet imagery (SVM), B= Planet imagery + NDVI (SVM), C = Planet imagery (DT), D = Planet imagery + NDVI (DT)) used as an input data in classification algorithms for crop mapping.

3.3.1.2 Decision Tree Algorithm

The decision tree algorithm was developed by (Deschamps et al., 2012). It has been widely used in classification studies due to its simplicity, flexibility, and computational efficiency (Watts & Lawrence, 2008). It creates multiple trees based on conditions to find the best way to classify a pixel into its corresponding class (Reddy, 2012). The final tree outcome is used as the best scenario found by DT. Several studies have investigated DT and reported its high classification efficiency (Khosravi et al., 2017; Verma, 2017; Liu et al., 2019). It has three parameters: the maximum number of trees, maximum tree depth, and the number of samples per class. The DT parameters were set by trial-and-error method (maximum number of trees = 50, maximum tree depth = 30,

and samples per class = 200) to achieve the optimal classification results on both data sets (i.e., multitemporal Planet imagery stand-alone (data set C), multitemporal Planet imagery in combination with multitemporal NDVI (data set D)) for crop mapping.

3.3.2 Accuracy Assessment and Performance Comparison of the Classification Algorithm

The literature describes several methods for determining the classification accuracy of maps (Stehman and Czaplewski, 1998; Smits et al., 1999). The error matrix, also known as the confusion matrix, is a better way to assess crop map classification accuracy (Foody, 2002). Producer accuracy, user accuracy, and the kappa coefficient are all included in the error matrix (Congalton, 2001). Correctly classified pixels indicated how the real feature on the ground surface was correctly shown on the classified map. The overall accuracy tells us how the overall algorithm performs in classification. The Kappa coefficient is another statistical indicator to evaluate classification accuracy. Kappa evaluates how well the classification has performed as compared to the randomly assigned value. Its values range from -1 to 1, with the lowest value indicating that the classification is not better than random classification. While a value close to a positive one indicates that the classification is significantly better than the random classification. The confusion matrix is not enough for classification performances; hence, the F1 score is an additional measure to provide the balance between producer and user accuracy (Chen et al., 2020). The performance of the algorithms was compared based on the overall accuracy and specific F1 score that took into account both producer and user accuracy (Rao et al., 2021).

$$\text{Producer/User Accuracy (\%)} = \frac{\text{Correctly classified pixels in one category}}{\text{Total classified pixels in all categories}} \times 100 \quad (3.1)$$

$$\text{Overall Accuracy (\%)} = \frac{\text{Number of correctly classified pixels in all categories}}{\text{Total number of referenced site pixels}} \times 100 \quad (3.2)$$

$$\text{Kappa Cofficient} = \frac{(TS - TCS) - \sum(Column\ total * Row\ Total)}{TS^2 - \sum(Column\ total * Row\ Total)} \times 100 \quad (3.3)$$

where *TS* is total samples, and *TCS* is total classified samples. Column sum and row sum mean the total classified pixels for each class in row and column.

$$F1\ score = 2\ \times \frac{Producer\ accuracy\ \times User\ accuracy}{Producer\ accuracy + User\ accuraccy} \quad (3.4)$$

3.4 Results

3.4.1 Normalized Difference Vegetation Index Multi-Temporal Profile for Each Class

Figure 3-2 shows the multitemporal NDVI profile for each class. According to the literature, the temporal profile results revealed that each crop has its distinct pattern during its various growth stages (Masialeti et al., 2010; Rao et al., 2021). Based on their growth cycles, the NDVI temporal profile helps discriminate between forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry (Figure 3-2). For example, the NDVI of potatoes increased rapidly from 0.43 to 0.77 during the productive stage, and then quickly declined from 0.75 to 0.4 during the maturity stage. At the peak of maturity, when the plant canopies were fully grown, potato crops had the greatest NDVI value. Wheat followed a similar pattern, although its NDVI value did not rise rapidly throughout the production stage. Soybean NDVI increased rapidly from May to mid-June, then dropped as the crop started to ripen and loosen its healthy green leaves. The NDVI profile of the urban class was completely different from that of all crop types (Figure 3-2), which helps to distinguish the urban class from crop types.

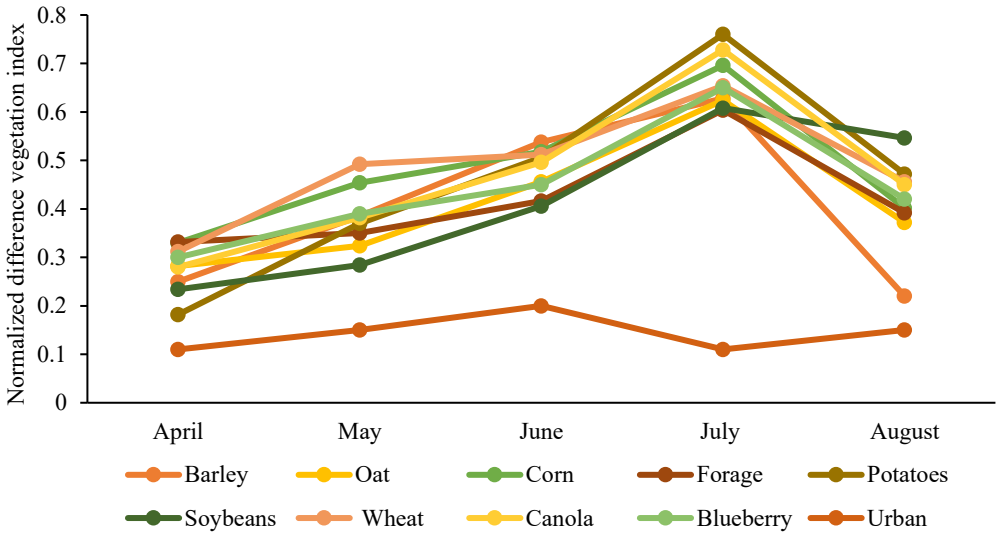


Figure 3-2. From growth to harvesting season, average monthly normalized differential vegetation index values for each crop type and urban class.

3.4.2 Performance of Support Vector Machine in Crop Mapping

The urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry were mapped using the input data sets (i.e., multitemporal Planet imagery stand-alone data set A, multitemporal Planet imagery combined with multitemporal NDVI data set B) in conjunction with SVM algorithm (Table 3-1). The SVM hit the acceptable user accuracy and producer accuracy for each class, namely, barley, blueberry, oat, canola, corn, forage, potatoes, soybeans, spring wheat, and urban of 72 to 90% when stand-alone multitemporal Planet imageries were utilized as an input data source (Table 3-2). Potato and forage achieved the highest user accuracy of 90%, while urban class yielded the lowest of 72% user accuracy (Table 3-2). The reason behind the lowest urban class user accuracy was pixels mixing with other crops. The F1 score yielded by urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry were 65, 93, 81, 85, 82, 89, 87, 79, 86, and 89, respectively (Table 3-2). The blueberry achieved the highest FI score, 93, and the lowest F1 score, 65, was achieved by urban (Table 3-2). The multitemporal Planet imagery

in conjunction with SVM-based classified maps achieved an overall accuracy of 83.8% and a kappa coefficient of 82% (Table 3-2). The calculated accuracy statistics are acceptable as a recently published study already reported the almost same accuracy range (Yang et al., 2019). Figure 3-3A represents the SVM-based classified map using a stand-alone multitemporal Planet imagery input data source.

Table 3-2. Producer accuracy, user accuracy, overall accuracy, kappa coefficient, and F1 score for crop mapping using multitemporal Planet imageries in conjunction with support vector machine algorithm.

Classes	Producer accuracy (%)	User accuracy (%)	Overall accuracy (%)	F1 score	Kappa coefficient
Urban	59	72	83.75	65	82
Forage	97	90		93	
Barely	79	85		81	
Oat	89	82		85	
Wheat	84	80		82	
Corn	92	87		89	
Canola	89	85		87	
Soybean	81	77		79	
Potato	82	90		86	
Blueberry	92	87		89	

Table 3-3. Producer accuracy, user accuracy, overall accuracy, kappa coefficient, and F1 score for crop mapping using multitemporal Planet imageries combined with multitemporal NDVI data in conjunction with support vector machine.

Classes	Producer accuracy (%)	User accuracy (%)	Overall accuracy (%)	F1 score	Kappa coefficient
Urban	79	85	90	81	89
Forage	90	90		90	
Barely	87	85		86	
Oat	94	90		92	
Wheat	95	92		93	
Corn	95	95		95	
Canola	93	92		92	
Soybean	92	87		89	

Potato	90	95		92
Blueberry	90	92		90

The producer accuracy, user accuracy, kappa coefficient, overall accuracy, and F1 increased significantly when multitemporal Planet imagery was combined with multitemporal NDVI data (data set B, Table 3-3) and used as the input data source in SVM. A 6.25% increase in overall accuracy and a 7% increase in kappa coefficient value were observed (Table 3-2,3). Onojeghuo et al. (2018) mapped rice fields, non-rice filed, built-up areas, and other vegetation and recorded that overall accuracy and kappa coefficient increased 8% and 11% after integrating the multitemporal NDVI data with Sentinel radar data. The user accuracy increased for urban, oat, spring wheat, corn, canola, soybeans, potato, and blueberry, 13, 8, 12, 8, 7, 15, 5, and 5 % respectively, when multitemporal NDVI data was combined with multitemporal Planet imagery (Tables 3-2,3). Forage and barely showed no increase in user accuracy when multitemporal NDVI data was fused with multitemporal Planet imagery (Table 3-3). Interestingly, no crop type class acquired less than 80% producer and user accuracy except urban class after fusing the multitemporal NDVI data with multitemporal Planet imagery (Table 3-3). The F1 score of each crop type also increased except forage when multitemporal Planet imagery fused with multitemporal NDVI was used in conjunction with SVM (Table 3-3). The compared result of SVM for both input data sets (i.e., multitemporal Planet imagery stand-alone data set A, multitemporal Planet imagery fused with multitemporal NDVI-data set B) proved that multitemporal Planet imagery fused with multitemporal NDVI is a better input data source for increasing the mapping accuracy and prepared map for this data set is represented in Figure 3-3B.

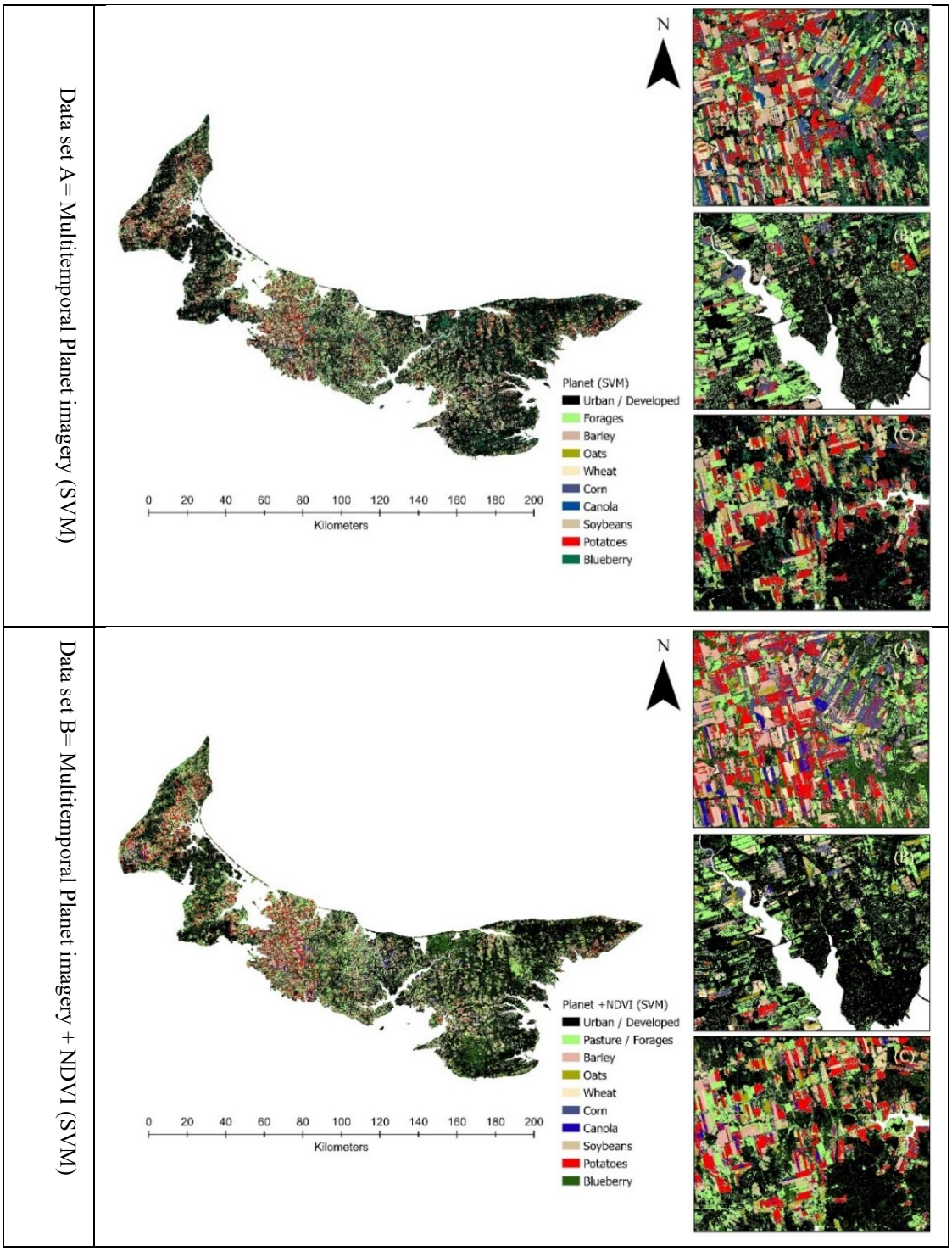


Figure 3-3. (A). Support vector machine-based classified maps using only multitemporal Planet imagery. (B). Support vector machine-based classified maps using multitemporal Planet imagery combined with multitemporal Planet imagery.

3.4.3 Performance of Decision Tree in Crop Mapping

The urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry were mapped using DT in conjunction with two different data sets (i.e., multitemporal Planet imagery stand-alone data set C, multitemporal Planet imagery combined with multitemporal NDVI data set D) (Table 3-1). The multitemporal Planet imagery stand-alone in conjunction with DT achieved user accuracy for each class between 75 to 90% (Table 4). The urban class had the lowest producer accuracy of 75%, and the oat class achieved the highest accuracy of 94% (Table 3-4). The overall accuracy and kappa coefficient achieved were 83 and 81% respectively (Table 3-4). The F1 score recorded for urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, and blueberry were 75, 81, 78, 80, 84, 90, 87, 84, 86, and 82, respectively (Table 3-4). The DT-based generated classified map using multitemporal Planet imagery as an input data source is presented in Figure 3-4A**.**

Table 3-4. Producer accuracy, user accuracy, overall accuracy, kappa coefficient, and F1 score for mapping using multitemporal Planet imagery in conjunction with the decision tree algorithm.

Classes	Producer accuracy (%)	User accuracy (%)	Overall accuracy (%)	F1 score	Kappa coefficient
Urban	75	75	83	75	81
Forage	80	82		81	
Barely	76	80		78	
Oat	83	77		80	
Wheat	86	82		84	
Corn	94	85		90	
Canola	85	90		87	
Soybean	87	82		84	
Potato	83	90		86	
Blueberry	81	83		82	

Table 3-5. Producer accuracy, user accuracy, overall accuracy, kappa coefficient, and F1 score for mapping using multitemporal Planet imagery combined with multitemporal NDVI in conjunction with decision tree algorithm.

Classes	Producer accuracy (%)	User accuracy (%)	Overall accuracy (%)	F1 score	Kappa coefficient
Urban	76	80	85	78	83
Forage	84	82		83	
Barely	78	90		84	
Oat	85	82		82	
Wheat	88	77		82	
Corn	94	85		89	
Canola	87	90		88	
Soybean	86	82		84	
Potato	88	92		90	
Blueberry	83	87		85	

The overall accuracy and kappa coefficient increased 2%, respectively, after adding the multitemporal NDVI data with multitemporal Planet imagery (Table 3-4,5, data set C vs D). The user accuracy increased for urban, barley, oat, canola, potato, and blueberry, 5, 10, 5, 5, 2, and 4 %, respectively (Table 3-4,5, data set C vs D). Interestingly, wheat observed a 5% decline; and forage and canola showed no change in the producer accuracy after blending the multitemporal NDVI with multitemporal Planet imagery (Table 3-4,5). The increase in the F1 score for urban, forage, barley, oat, spring wheat, corn, potato, and blueberry was 3, 2, 6, 2, 2, 1, 4, 3, respectively, when Planet imagery combined with multitemporal NDVI data (Table 3-4,5). The corn recorded a decline in the F1 score, and soybean showed the same F1 score (Table 3-4,5, data set C vs D). The calculated accuracy statistics documented that mapping accuracy increased when multitemporal NDVI data was combined with multitemporal Planet imagery. The DT-based classified map using the multitemporal Planet imagery combined with multitemporal NDVI as an input data source is presented in Figure 3-4B.

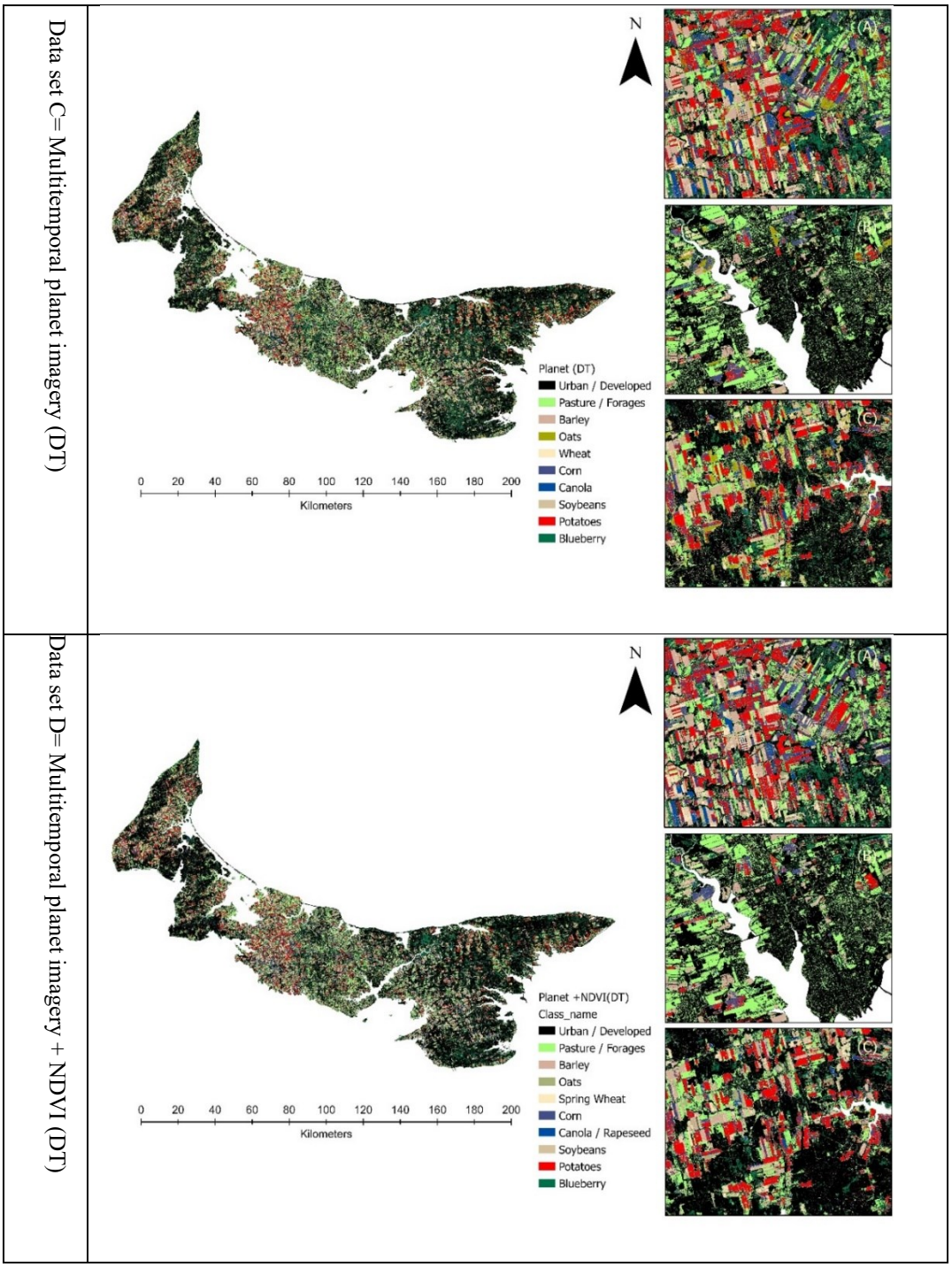


Figure 3-4. (A). Decision Tree-based classified maps using only multitemporal Planet imagery. (B). Decision Tree-based classified maps using multitemporal Planet imagery combined with multitemporal Planet imagery.

3.4.4 Comparative Performance of DT and SVM for Optimal Output

Based on the classification results of SVM and DT in sections 4.2 and 4.3, the most promising result was produced by SVM and DT when the combined multitemporal Planet imagery with multitemporal NDVI was used as an input data source (Table 3-1, data set B, and D). The SVM classification produced an overall accuracy 5% higher than the DT-based classification (Table 3-3,5). SVM-based producer accuracy was higher than the DT-based producer accuracy (Tables 3-3,5). Interestingly, the kappa coefficient value of SVM-based classification was 6% higher than the DT-based classification using the same input data source (i.e., multitemporal Planet imagery combined with multitemporal NDVI). SVM-based classification yielded higher FI scores for each class than the DT-based classification (Figure 3-5). SVM-based classification observed the 3, 7, 2, 10, 11, 6, 4, 5, 2, and 5 points increase for urban, forage, barley, oat, spring wheat, corn, canola, soybeans, potato, blueberry, respectively as compared to the DT-based F1 scores (Table 3-3,5). The optimal classifier for crop mapping was decided by comparing the overall accuracy and F1 scores (Rao et al., 2021). The calculated accuracy statistics proved that classified maps produced with SVM achieved higher accuracy than DT-based classified maps no matter which input data source was used (Table 3-2,5). This result is also confirmed by already published research; for example, Vibhute and Gawali (2013) indicated that SVM has higher accuracy for crop type mapping than RF. Kumar et al. (2017) used the SVM and maximum likelihood classifier to map the crops, namely, wheat, lentils, mustard, pigeon pea, linseed, corn, and sugar cane and reported that SVM algorithms performed better than the maximum likelihood classifier.

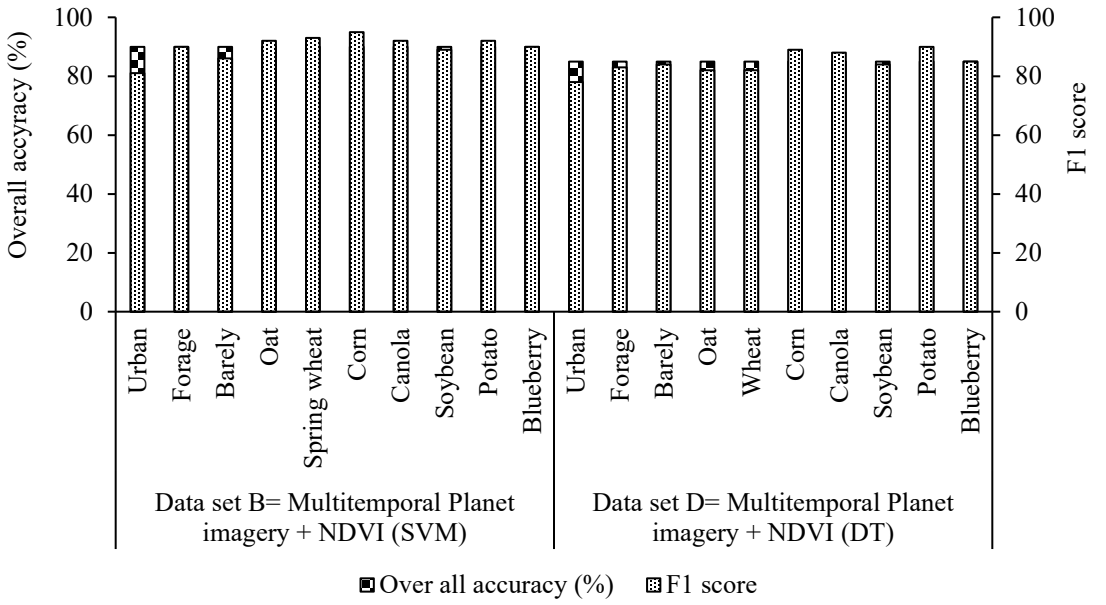


Figure 3-5. The overall accuracy and F1 scores comparison between Support Vector Machine and Decision Tree for each classified class using the combined multitemporal Planet imagery with multitemporal NDVI data as an input data source.

3.5 Discussion

This study explored the potential of multitemporal Planet imagery alone and multitemporal Planet imagery integrated with multitemporal NDVI for preparing crop maps at a broad scale using a machine learning algorithm. The crop maps of PEI were created at 3 m resolution using the SVM and DT machine learning classification algorithms on four data sets (i.e., A= Planet imagery (SVM), B= Planet imagery + NDVI (SVM), C = Planet imagery (DT), D = Planet imagery + NDVI (DT)). Using the multitemporal Planet imagery stand-alone in conjunction with SVM and DT achieved almost the same overall accuracy of 83.75% and 83%, respectively (Table 3-2,4). Interestingly, the trend remained the same: the highest user accuracy was achieved by potato and the lowest user accuracy was yielded by urban class in all data sets (Table 3-2,5,6,7). These results

are better than those from studies that mapped the crop from different data sources; for example, Recently, Rao et al. (2021) documented that Planet imagery in conjunction with SVM yielded overall mapping accuracy of 0.73%. Ponya et al. (2020) used the time series Sentinel-2 data for crop mapping using SVM and DT and achieved an overall accuracy of about 73%.

In addition to the multitemporal Planet imagery, the inclusion of optical derived vegetation indices such as NDVI depicting the crop phenology information can help to provide more accurate methods to differentiate the crop classed from other land cover classes (Onojeghuo et al., 2018). This experiment also assessed the performance of two machine learning algorithms (i.e., SVM, and DT) when the combined multitemporal Planet imagery with multitemporal NDVI was used as an input data source. A previous study reported that the addition of multitemporal vegetation indices data in multitemporal satellite imagery facilitates improving classification accuracy (Rajah et al., 2019). The accuracy statistics supported the hypothesis of previous studies that the SVM and DT algorithms can result in a significant increase in the overall accuracy when multitemporal NDVI data is combined with Planet imagery (Tables 3-5,7). The results indicated that accuracy increased by combining the vegetation indices with satellite data, but a little part of accuracy depends on the classifier's performance.

The result of the 3rd step of this analysis proved that SVM is a better algorithm than DT algorithms. The SVM algorithm outperformed and yielded a 5% higher overall accuracy as compared to the DT algorithm (Tables 3-3,5). The results proved that SVM is the best optimal algorithm for crop mapping, whether multitemporal NDVI data is combined with planet imagery or not. All the results of this study also proved that the SVM algorithm outperformed than DT algorithm, without concerning the input data source (Tables 3-2,3,4,5). The documented result of

this study is similar to previous studies for example; Devadas et al. (2012) found that SVM is more accurate in classification as compared to the maximum likelihood classifier.

Agriculture and Agri-Food Canada prepared the crop maps for PEI at 30 m spatial resolution using DT and achieved an overall accuracy of 85% (Illert & Afflerbach 2020). In this study, SVM achieved 90% overall accuracy when multitemporal Planet imagery was combined with multitemporal NDVI and used as an input data source, which is 5% higher than AAFC published maps. The three main reasons for the higher accuracy of these maps are i) higher spatial resolution, ii) integrated original satellite data with NDVI data used in classification, and iii) selection of the appropriate classification algorithm.

3.6 Conclusion

The results show that using combined multitemporal Planet imagery and multitemporal NDVI data source as an input improved the mapping accuracy of the SVM and DT algorithms. The SVM, on the other hand, was shown to be the best method for mapping without relying on the input data sets. When multitemporal Planet imagery was combined with multitemporal NDVI as an input data source, SVM achieved more accurate results. The overall accuracy and kappa coefficient for this method were 90 and 80 %, respectively. All of our findings revealed that combining data sources can improve mapping accuracy (i.e., multitemporal Planet imagery combined with multitemporal NDVI). These findings can aid in the creation of higher-accuracy and higher-resolution crop maps for any location, which can aid in the development of agriculture policy, acreage estimation, and decision-making. To highlight crop phonology, several vegetation indexes have been produced. By integrating diverse vegetation with Planet imagery, crop mapping accuracy may be evaluated in the future.

CHAPTER 4: SATELLITE-BASED CROP EVAPOTRANSPIRATION MAPPING TO ACCESS THE VARIATION CONCERNING THE CROP DEVELOPMENT

ABSTRACT

Water for irrigation is in scarce supply in numerous regions around the world. Water management is greatly influenced by the accurate calculation of crop evapotranspiration (ET_c). The use of remote sensing to estimate ET_c is a quick and reliable method. This work shows how remote sensing data and GIS can be used to create NDVI-based potato crop coefficient (K_c), ET_c maps at 3 m resolution for sustainable water management. The four Planet images were downloaded to cover four potato growth stages, i.e., (initial, development, mid, and late) during the 2021 season over the designated potato field ($46.24136° N, 63.5311° W$) located in Prince county in Prince Edward Island (PEI), Canada. A higher correlation $R^2 = 0.95$ was found between Food and Agriculture Organization (FAO) suggested K_c and NDVI. The developed regression linear equation was used as input in GIS for developing the new K_c of the potato during the four potato growth development stages. The Hargreaves method was used to estimate the reference evapotranspiration (ET_r) corresponding to four potato growth stages. The final potato ET_c maps were prepared by multiplying the ET_r with newly developed K_c maps. The result indicated that the potato ET_c value was low at the initial stage of growth from 0.69 to 2.01 mm/day, and increased from 1.89 to 3.05 mm/day, at the development stage of potato growth. The highest ET_c value was noted at mid growth stage from 2.02 to 4.02 mm/day; then, a decline was observed in ET_c values at the late stage from 1.06 to 2.16 mm/day. The developed ET_c maps assist in estimating the actual potato water requirement variability in a field over four growth stages. By using the developed

ET_c maps, farmers can apply the proper amount of water to fulfill the actual potato water requirements and lead to water conservation.

Keywords: Remote-sensing; Geographic information system; Planet imagery; Normalized difference vegetation index; Hargraves method; Reference evapotranspiration; Crop evapotranspiration

4.1 Introduction

Potato is considered the fourth most important food crop and the primary agricultural commodity of PEI. The potato industry contributes $1 billion annually to the economy of PEI (Macdonald, 2020), roughly 10.8% of the provincial gross domestic product (GDP), and 12% of the jobs for the local islanders each year (Agriculture and Agri-Food Canada, 2017). Potato yield is very sensitive under limited water conditions (Food and Agriculture Organization, 2021). The potato production in PEI is rainfed therefore, maximum rainfall is required to achieve a higher potato yield (Afzaal et al., 2020). However, the climate of PEI is changing now, facing more hot days and fewer cold nights, and precipitation patterns are also changing (Maqsood et al., 2020). To mitigate this precipitation pattern-changing problem, supplemental irrigation is needed to fulfill the potato water requirement to stay competitive in the global market (Afzaal et al., 2020). In 2018, the average potato yield in Alberta decreased to 272 cwt per acre. This yield was increased to 392 cwt per acre due to supplemental irrigation (Belova, 2018). Shock et al. (1998) reported the same trend as potato yield decreased in a water shortage.

The provincial government of PEI has placed a moratorium on well drilling for agricultural irrigations. The farmers who had wells before the moratorium can now only pump the water from the wells, store it in ponds, and later use it for irrigation (Belova, 2018). As a result, irrigation water is scarce in PEI, and better estimation of crop water requirements is essential to use the water efficiently. For water conservation, farmers must adopt new methodologies for estimating the crop water requirement accurately.

A prominent method to estimate ET_c is to multiply the ET_r by K_c (Adamala et al., 2016). The penman monteith equation is considered a standard equation to estimate the ET_r, however, it requires a lot of metrological parameters (e.g., solar radiation, air temperature, wind speed) (Allen

et al., 2005). A recently published study discussed the potential of the Hargreaves method to estimate the ET_r with limited metrological parameters (Maqsood et al., 2020). Typically, the K_c value is taken from the literature corresponding to crop type and its growth stages (Reyes-Gonzalez et al., 2018; Reyes-Gonzalez et al., 2015). The crop evapotranspiration is estimated using different point-based methods (e.g., weighing lysimeter, evaporation pan, Atmometer) but the maintenance cost is high for these methods (Xiang et al., 2020).

In this era, remote-sensing is a viable technology to estimate ET_c in less cost and time (Kjaersgaard et al., 2011). Two methods, empirical and physical-based, have been developed to estimate the evapotranspiration from remote-sensing data (Allen, R. G. et al., 1998; Lei & Yang, 2014). The empirical method uses time-series indices and metrological data, while the physical method uses the surface energy balance and latent heat flux to estimate ET_c (Glenn et al., 2007). In the empirical method of remote sensing, the K_c can be calculated based on the spectral reflectance of vegetation indices. The preparation of ET_c maps based on remote sensing is an alternative to the point-based methods.

Lei & Yang (2014) proved that vegetation indices could estimate the K_c of different crops as vegetation indices showed a higher correlation with Food and Agriculture Organization (FAO) K_c value for different crops. The NDVI is the most common index out of different vegetation indices used to estimate K_c of different crops (Glenn et al., 2011). For example, the K_c depend on the NDVI derived from Sentinel images at 10 m spatial resolution was developed by establishing the relationship between sugarcane K_c values taken from the FAO-56 manual. From this relationship, an R^2 value of 0.70 was achieved (Alface et al., 2019). Farg et al. (2012) developed the K_c for the wheat crop from NDVI derived from the SPOT-4 satellite and achieved the R^2 0.82, 0.92, and 0.97 during the first, second, and third growth stages respectively. One study reported a

high correlation (R^2 = 0.72) between soybean K_c values derived from NDVI and the suggested value of K_c by FAO-56 (de Oliveira et al., 2016). Landsat-8 imagery was used to establish the relationship between NDVI, and K_c values suggested by the FA0-56 manual of groundnut and an R^2 value of 0.75 was achieved. The higher R^2 = 0.71 was achieved by developing the relationship between wheat K_c value taken from the FAO-56 manual and NDVI derived from the Indian Remote Sensing satellite (Gontia & Tiwari, 2010). To the best of my knowledge, no relationship exists between the Planet satellite-based derived NDVI and K_c values taken from the FAO-56 manual for developing the NDVI based potato K_c in the Canadian eastern province PEI.

In this study, a relationship between NDVI derived from the Planet imagery at a higher resolution of 3 m and the K_c value taken from the FAO-56 manual for a potato field in the prince county of PEI was established. This established relationship will be used to prepare the potato K_c maps. The Hargreaves method was selected due to limited climatic parameters availability to calculate the ET_r and the calculated ET_r was used to make the Potato ET_c maps at 3 m resolution. No studies have been done to estimate the potato ET_c maps in PEI by using Planet satellite imagery.

Water is scarce in PEI for irrigation purposes and climate change demands supplemental irrigation for better potato yield. The developed ET_c maps at 3 m resolution will help estimate the spatial and temporal variability of potato crop water need at the field scale. This developed potato ET_c maps at four growth stages (i.e., initial, mid, development and late) will help the farmers to apply the water precisely and help them to save the water for future use.

4.2 Material and Methods

4.2.1 Study Area

The study was carried out on a selected potato field ($46.24136° N, 63.5311° W$) during the growing season of 2021 in PEI, shown in Figure 4-1. The common potato variety Russet Burbank was cultivated in this field. In terms of management practices, this field was cultivated under the conventional management practices of the region. Fertility and pest management were applied based on the provincial recommendation. The distance between ridges was 0.75 m, and furrow depth was about 0.15 to 2.0 m. In this study area, the potato crops are sown from mid-March to the end of April and the harvesting session spans from mid-September to late October.

Figure 4-1. Zoomed location of study field (red colour) in Prince Edward Island, Canada.

4.2.2 Planet Satellite Data

The Planet satellite takes the earth images at 3 m spatial resolution. It has four spectral bands (Blue: 455- 515 nm; Green: 500-590 nm; Red: 590-670 nm; and Near Infrared (NIR): 780-860 nm). The revisit time of this satellite is about one day to cover 340 million km^2/day (Planet Team, 2021). The four satellite images were downloaded to cover four potato growth stages (i.e., initial

stage, development stage, mid stage, and late stage). The cloud-free images were selected to avoid the absorption and scattering problem, then K_c depending on the NDVI was developed.

4.2.3 NDVI Calculation and Pixel Selection Criteria for Average NDVI Value

Satellite sensor measures red and NIR light waves reflected by the land surface (Vladimirova et al., 2008). The mathematical algorithms change this raw satellite data into vegetation indices (Bannari et al., 1995). The vegetation indices are indicator measures of the greenness and health of vegetation (Muraoka et al., 2013). Several vegetation indices are available, but the NDVI is the most widely used (Matese & Di Gennaro, 2021). The NDVI values were calculated during the initial, development, mid and late potato crop growth stages with the help of ArcGIS Pro software and by using this formula $NIR - Red/NIR + Red$. After calculating the NDVI for the potato field, 10 pixels were randomly selected to calculate the average NDVI value during each growth stage. We assumed that these selected pixels are true representatives of the entire potato field. The same pixel values were observed throughout the potato growth stages.

4.2.4 Relationship Establishment Between NDVI and K_c Maps Development

The K_c depends on the crop growth stages and variates during different growth stages. Potato growth stages are divided into initial (20 days after planting (DAP)), development stage (21-50 DAP), mid (51-110 DAP) and late-stage (110-140). The K_c value fluctuates according to these growth stages: for example, the K_c value varies from 0.4 to 0.5 at the initial stage, between 0.7 to 0.8 in the development stage, from 1.05 to 1.2 in the mid-stage of growth, and from 0.7 to 0.75 in late session (Food and agriculture organization, 2021).

We have adjusted the k_c value according to these stages and developed the relationship between NDVI and K_c. This established relationship was used to generate the linear regression equation.

The established linear regression equation between NDVI and K_c help us to develop the K_c maps using ArcGIS Pro software.

4.3 Reference Evapotranspiration Calculation

4.3.1 Metrological Data Acquisition

An affordable and accurate weather monitoring system is essentially needed in small-scale research such as precision agriculture (Dombrowski et al., 2021). A variety of weather stations are available in the market, but the low cost and non-standard technologies raise concerns about data quality. A stand-alone metrological station was installed on the field (longitude $-63.84°$, and latitude $46.44°$) to access the maximum and minimum temperature corresponding to the acquired satellite images date. The time series readings of the metrological parameters were downloaded from the datalogger of the weather stations.

4.3.2 Hargreaves Method to Estimate the Reference Evapotranspiration

The penman-monteith is a standard method to calculate the ET_r, but it needs many metrological parameters such as temperature, wind, velocity, solar radiations, and moisture content (Cai et al., 2007). In recent publications, the Hargreaves method was used to estimate the ET_r, as this method requires fewer metrological parameters (Allen, 2001; Rodrigues and Braga, 2021). The required parameters for the Hargreaves method include the maximum and minimum temperature and extraterrestrial radiation. Extraterrestrial radiation is calculated based on the day of the year and latitude (Patel et al., 2015). The Hargreaves method was selected to estimate potato crop ET_r in PEI due to its simplicity and acceptable accuracy results. The Hargreaves equation used in this study is:

$$ET_r = 0.408(0.0023)(17.8 + Tmean)(Tmax - Tmin)^{0.5}\, Ra \quad \text{.....(4.1)}$$

Ra is extraterrestrial radiations (MJm^{-2}) and 0.408 is an empirical factor to convert the MJm^{-2} to mm.

The solar radiation at the top of the earth's atmosphere is called Ra. The given equation is calculated based on the location's latitude and calendar day of the year.

$$Ra = \frac{24(60)}{\pi} \times G_{sc}(d_r)[(\sin(\varphi) \times \omega_s \times \sin(\delta)) + (\cos(\varphi) \times \cos(\omega_s) \times \cos(\delta))] \ldots\ldots\ldots\ldots (4.2)$$

G_{SC} = The solar constant (0.0820 MJm^{-2})

φ = Location in radian.

d_r = The inverse relative distance from the earth to the sun and calculated by using this equation:

$$d_r = 1 + [0.33 \times \cos(\frac{2\pi}{360} J)] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.3)$$

Where J represents the day of the year.

ω_s = Sunset hours in radians and calculated by using this equation:

$$\omega_s = arccos(-\tan(\varphi) \times \tan(\delta)) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.4)$$

Where δ represents the solar declination in radians and is calculated as:

$$\delta = 0.409 \times \sin\left(\frac{2\pi}{360} \times J\right) - 1.39 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.5)$$

4.3.3 Potato ET_c Map Development

The K_c maps multiplied with calculated ET_r to generate the ET_c maps of potato crop at a higher spatial resolution of 3 m, for initial, development, mid and late growth stages using the ArcGIS

Pro software. The developed ET_c maps will be used to evaluate the water requirement of potatoes during four growth stages such as initial, development, mid, and late.

4.3.4 Flow Chart of this Whole Study

The entire procedure to prepare a potato ET_c map based on Planet satellite vegetation indices is represented in Figures 4-2. The Planet satellite-based calculated NDVI, vegetation indices-based K_c, and metrological data are the major data inputs for preparing the potato ET_c maps during four growth stages.

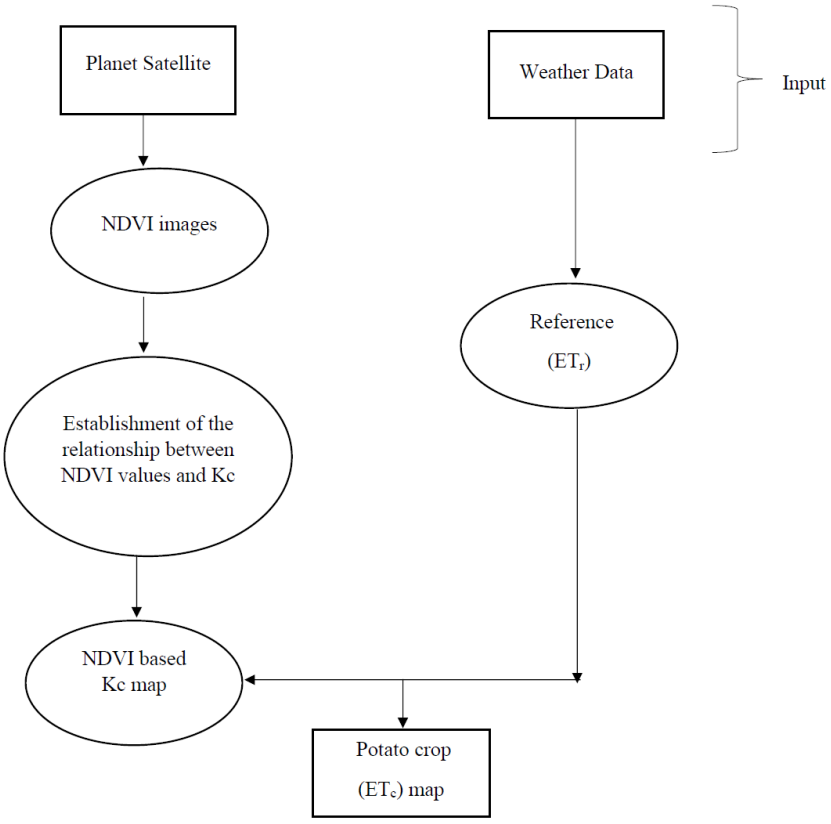
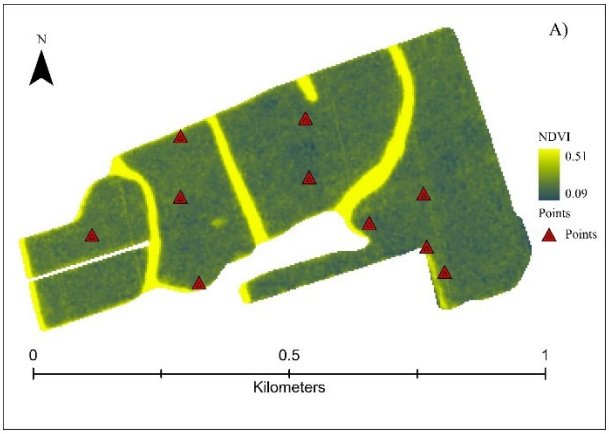


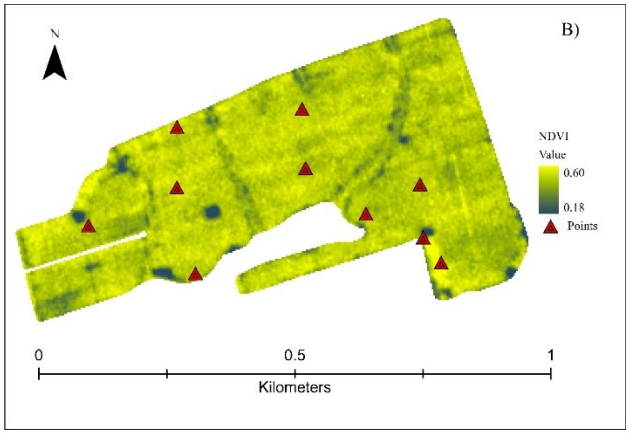
Figure 4-2. The flow diagram for explaining the whole process to develop the ET_c maps of the potato crop.

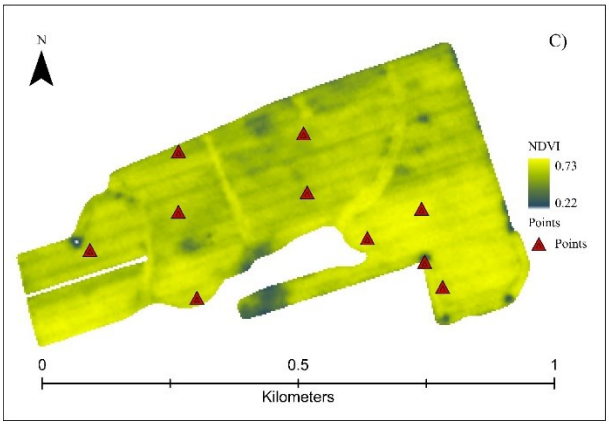
4.4 Result and Discussion

4.4.1 Normalized Difference Vegetation Index Relationship With FAO Suggested K_c

The NDVI average values extracted from the selected pixels in the potato field and growth stages are represented in Figure 4-2. In the initial stage of development, the NDVI values were low, around 0.11 (Figure 4-3A). During the development stage, the values of NDVI increased (0.11 to 0.49), and NDVI value increased (0.49 to 0.69) during the late growth (Figure 4-3B, C). A decline was observed in NDVI value during the final stage (Figure 4-3D). Adamala et al. (2016) assessed the NDVI values for wheat and demonstrated the same pattern, such as low NDVI value at the initial stage and high NDVI value at the development stage. For this selected potato field, a strong relationship between NDVI and FAO K_c was observed (R^2=0.88) and a linear regression equation $K_c = 1.043NDVI + 0.4159$ was established (Figure 4).







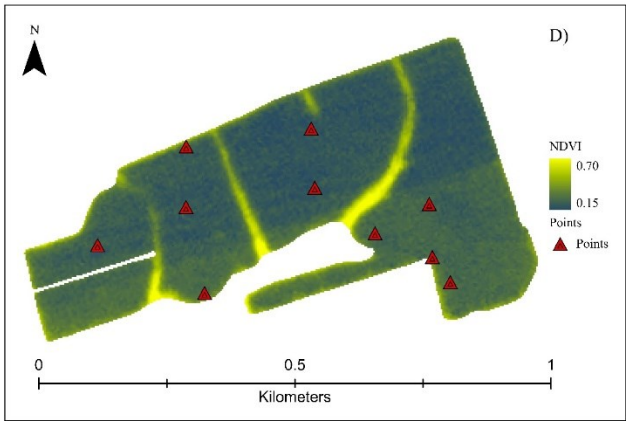


Figure 4-3. A) Selected pixels for NDVI value estimation in a potato field at the initial stage of growth. B) Selected pixels for NDVI value estimation in a potato field at the development stage of growth. C) Selected pixels for NDVI value estimation in a potato field at the mid-stage of growth. D) Selected pixels for NDVI value estimation in a potato field at the late stage of growth.

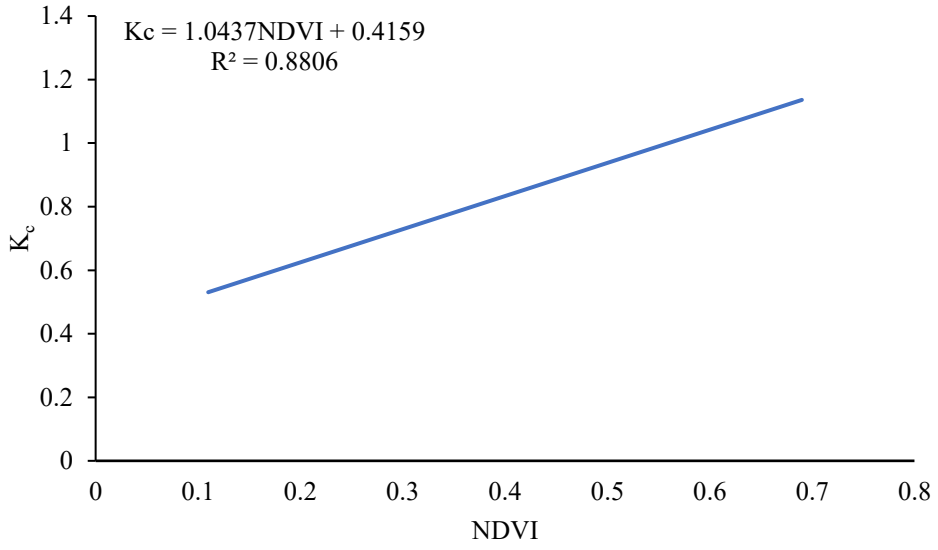


Figure 4-4. The established relationship between FAO suggested potato K_c and Planet imagery-based NDVI values during the 2021 growing session.

4.4.2 NDVI Based K_c Maps

The established linear equation between NDVI and FAO potato K_c was used to generate the K_c maps at 3 m resolution. Figure 4-5 shows the temporal and spatial variation of K_c values in the selected potato field during the growing season of 2021. At the initial stage, the developed K_c maps showed low Kc values of 0.32 to 0.5 (cool green to light green), which is very close to potato K_c values suggested by FAO (Figure 4-5A). The Kc map prepared based on NDVI showed the highest patches of K_c values in a field where some common grasses and other vegetation were present (Figure 4-5A). The K_c values increased at the development stage of the potato crop and varied

from 0.61 to 0.85 throughout the field (Figure 4-5B), and at this stage, FAO suggested the potato K_c value was 0.8 which is very close to NDVI based K_c value. At the mid-stage, the K_c maps showed a value from 0.65 to 0.75 (cool green to yellowish-green) (Figure 4-5C). At the late stage of development, the K_c value varies from 0.57 to 1.14 (Figure 4-5D). Lei & Yang (2014) developed the K_c maps based on NDVI calculated from Landsat imagery for winter wheat and summer maize to understand better the spatial and temporal distribution of K_c value in a field.

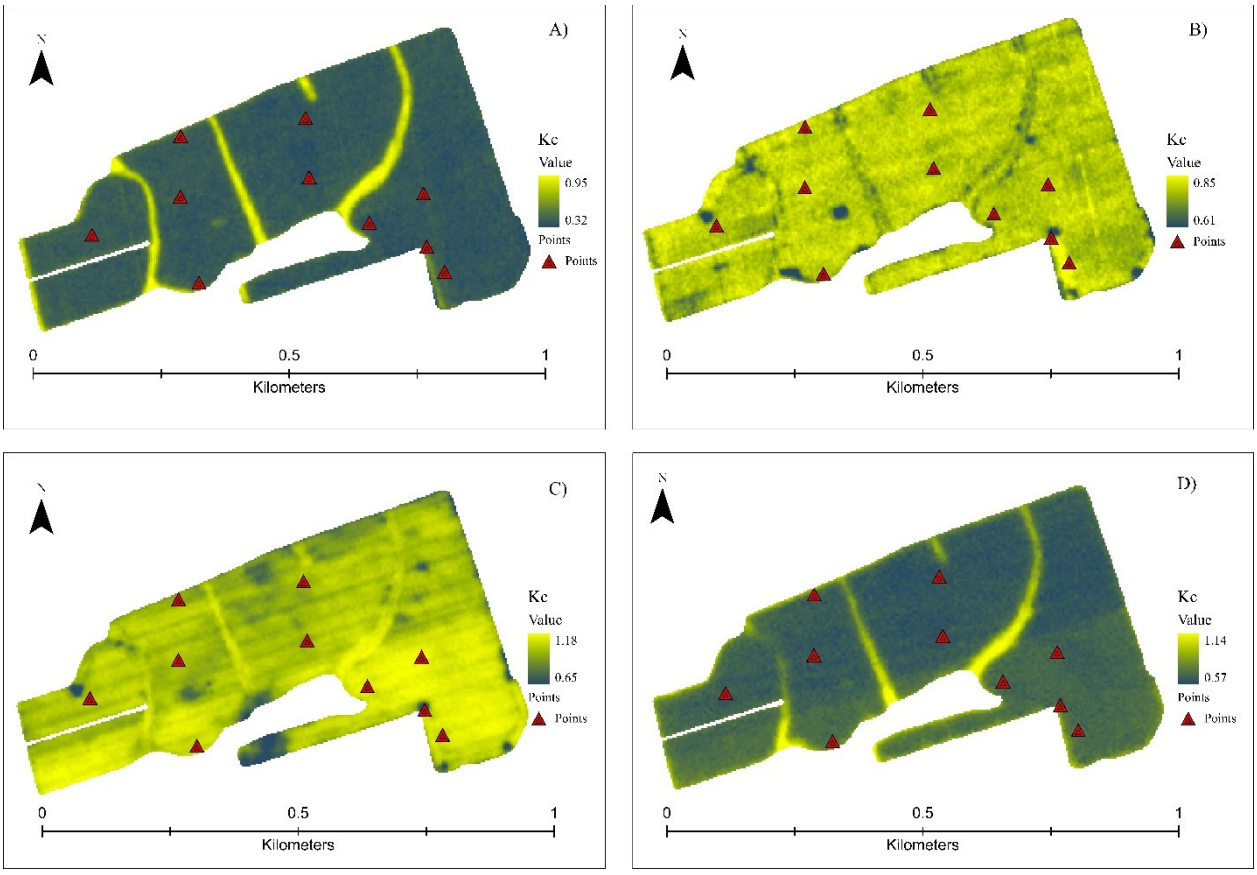


Figure 4-5. A) Spatial and temporal variability of potato K_c value at the initial stage. B) Spatial and temporal variability of potato K_c value at the development stage. C) Spatial and temporal variability of potato K_c value at the mid-stage. D) Spatial and temporal variability of potato K_c value at a late stage.

The relationship between FAO suggested potato K_c and NDVI-based calculated K_c was shown in Figure 4-6. A strong relationship was founded ($R^2 = 0.95$). This relationship proved that the K_c values developed based on the NDVI are robust parameters to estimate the actual potato ET_c at field level.

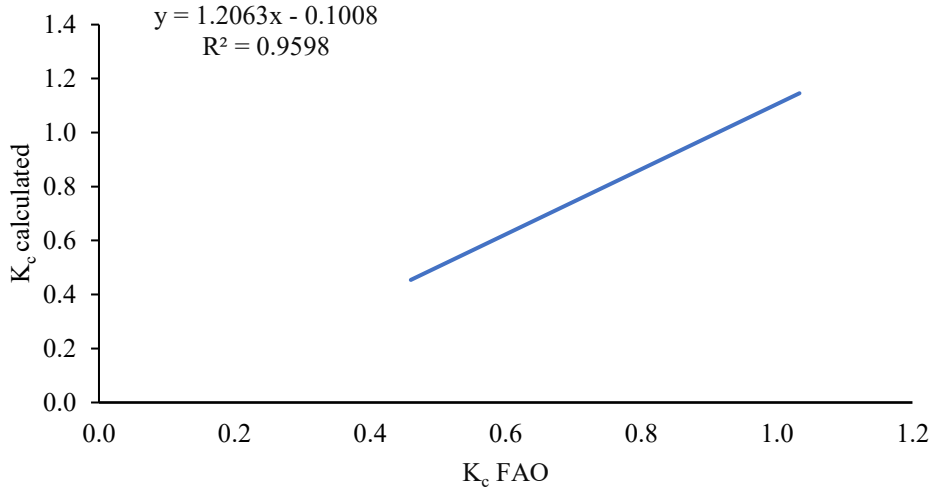


Figure 4-6. Relationship between K_c calculated, and FAO suggested K_c for potato during the growing season of 2021.

4.4.3 Potato Crop ET_c Maps

The potato ET_c maps at 3 m resolution (Figure 4-7A, D) were developed by multiplying the NDVI-based Kc with the Hargrave method-based calculated ET_r. The ET_c map showed a low ET_c value of 0.69 to 2.05 mm/day (green to red colour) at the initial stage of growth (Figure 4-7A). The highest value of ET_c showed in the patches because some common grasses and other vegetation covered these patches. At the development stage, the potato ET_c showed a high value of ET_c from 1.89 to 3.15 mm/day (green to red colour) (Figure 4-7B). The highest value of ET_c was observed in the mid-stage of growth from 2.22 to 4.02 mm/day (green to red colour) (Figure

4-7C). At the late stage of development, ET_c values declined throughout the field (Figure 4-7D). Gontia & Tiwari (2010) also supports this study's results; for example, the initial stage crop needs lesser amounts of water mean less ET_c, and as the crop grows, the water requirement increases means the ET_c value increases. Most of this research prepared ET_c maps based on NDVI for different crops, but they showed the same trend as shown in our study (Adamala et al., 2016; French et al., 2020). Simonneaux et al. (2008) documented that remote sensing-based ET_c estimation is a promising method. Potato ET_c maps prepared at a field-scale level at a higher 3 m resolution are helpful to quantify the actual water requirement corresponding to its different development stages. These ET_c maps are also helpful for the PEI farmers to design the irrigation schedule of their potato crop to achieve a higher yield and enhance irrigation water management.

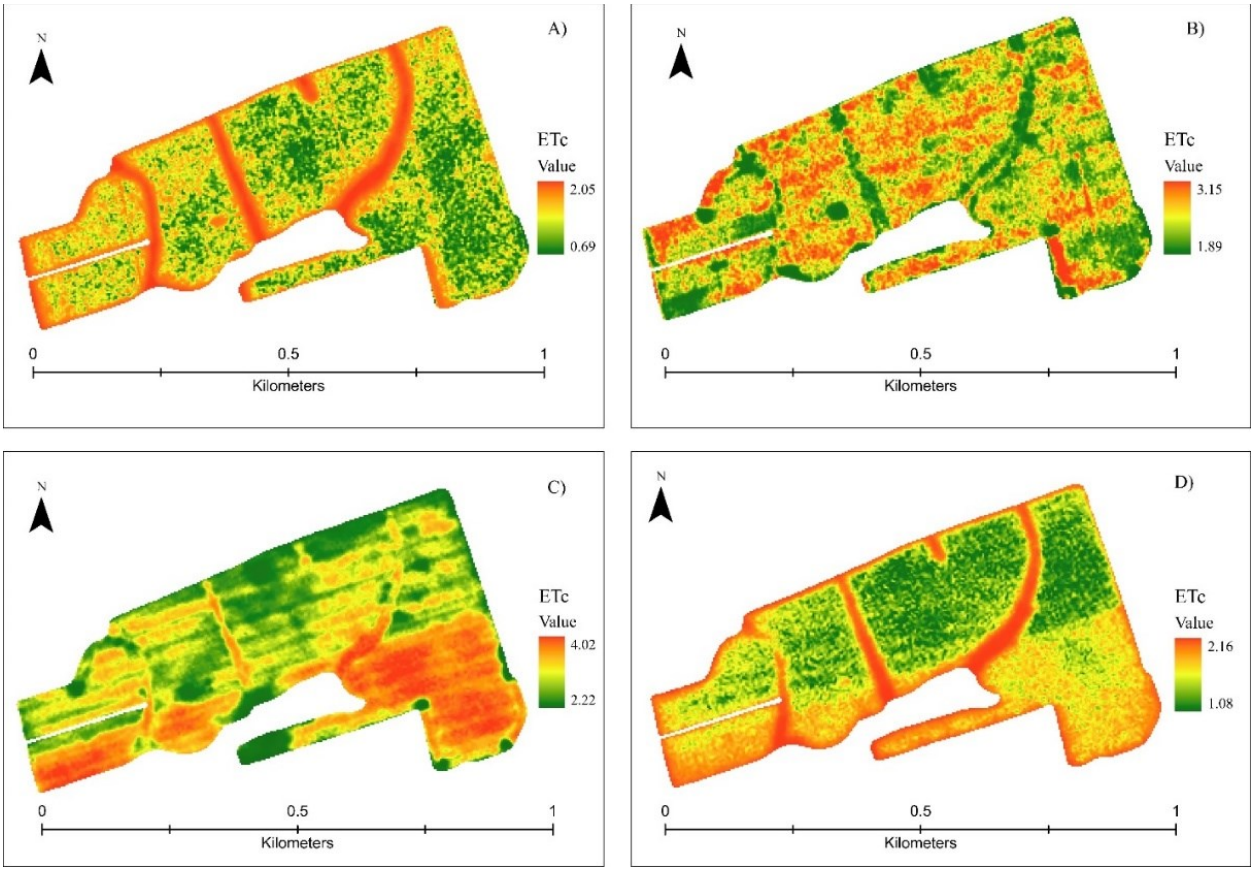


Figure 4-7. A) Spatial and temporal variability of potato ET_c value at the initial stage. B) Spatial and temporal variability of potato ET_c value at the development stage. C) Spatial and temporal variability of potato ET_c value at the mid-stage. D) Spatial and temporal variability of potato ET_c value at a late stage.

4.5 Discussion

This study presented a planet satellite-based approach for preparing NDVI-based ET_c maps of the potato crop. The limitation of climatic variables became a strong reason for choosing the Hargraves method for the estimation of ET_r for potato crops at four different growth stages. Valipour and Eslamian (2014) reported that the Hargraves method is suitable for estimation of ET_r as compared to the other temperature-based methods, i.e., Modified Baier Robertson, Modified Hargreaves-Samani 1, Modified Hargreaves-Samani 2, Modified Hargreaves-Samani 3, and Modified Hargreaves-Samani 4. The selected Hargreaves method was used to estimate the ET_r for potato crops due to the limited availability of climatic variables.

The FAO-suggested K_c values for potatoes showed a higher correlation $R^2 = 0.95$ with NDVI-based calculated K_c values of potatoes (Figure 4-6). The developed Kc values showed no difference with FAO-suggested K_c values at the initial and development potato growth stage, but a higher K_c value was observed at the development stage of the potato (Figure 4-8). At a late stage, no difference was observed between FAO-suggested K_c and NDVI-based K_c (Figure 4-8). Reyes-González et al. (2018) reported the correlation $R^2 = 0.93$ between NDVI-based K_c and FAO-based K_c for corn. French et al. (2020) used the NDVI to develop a new K_c for durum wheat in the US southwest for ET_c estimation. Zhang et al. (2019) compared the K_c value calculated on the field for maize with FAO K_c values and found that K_c developed based on vegetation indices has greater

potential for estimating the maize ET_c. This literature supported the selection of the NDVI vegetation indices for establishing the relationship with FAO-suggested K_c for developing the new potato K_c based on NDVI. The newly developed potato K_c helped to estimate the actual water requirement corresponding to potato growth stages.

The variability in the ET_c value was identified by visualizing the prepared ET_c maps during each growth stage (Figure 4-7A, D). In Figure 4-7A, potato ET_c values have varied from 0.69 to 2.05 mm/day (green to red colour) in the field, the higher variation in the ET_c value depends upon multiple factors i.e., soil moisture content, shade cover, temperature (Lin 2010; Nistor et al., 2016). In the initial stage of growth, the ET_c was less but grass patches in the field (Figure 4-7) became a strong reason for higher variation in the ET_c value. In the development and mid-stage, attained the value of ET_c 1.89 to 3.05 and 2.22 to 4.02 mm/day respectively for potato

Aghajanloo et al. (2013) used the ANN genetic algorithm to estimate the potato ET_c in Iran and found the mean and maximum potato ET_c were 0.49 and 4.16 mm/day, which is very close to the present results. In the late stage of growth, the ET_c value declined because potato crop water consumption decreased. The trend was compared with accepted literature to support the result of this study. For example, Tasumi & Allen (2007) and Morais et al. (2015) document that ET_c values change due to the timing of crop development.

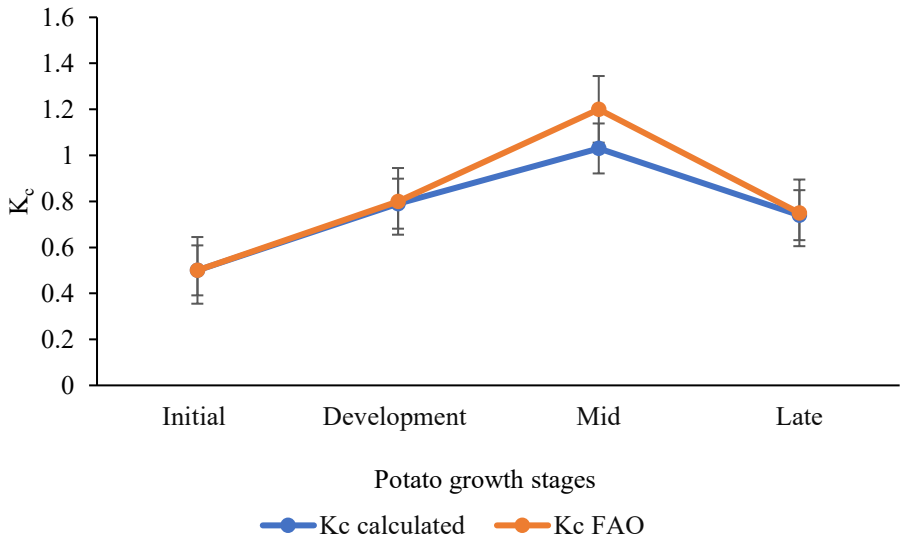


Figure 4-8. The comparison between potato FAO K_c and NDVI-based potato K_c during four growth stages.

4.6 Conclusion

The core objective of this study was to prepare ET_c maps of potato crops based on the Planet satellite imagery for a selected field in PEI, Canada. A linear regression equation was developed to prepare new K_c values during the four potato growth stages. The generated linear regression was used in ArcGIS Pro as input to develop the K_c maps. The developed ET_c value is low at the initial stage of growth, high at the mid-stage of growth, and again declines at the harvest stage of growth. The developed ET_c value is low in the early stages of development, rises in the middle stages, and then falls in the harvest stage. These maps aid in quantifying the variability of potato crop water requirements at the field size. This study might be expanded in the future to include the use of generated ET_c maps of potato crops for long-term irrigation water management.

CHAPTER 5: CONCLUSION

This study evaluated the potential of satellite-based indices for 1) preparing LULC maps, 2) improving crop mapping accuracy, and 3) preparing potato ET_c maps in conjunction with machine learning algorithms. Three separate studies were completed to achieve these specific objectives. In study 1, the agriculture, urban, barren land, bare soil, and forest were mapped using the Landsat-8 and Sentinel-2A data. In conjunction with RF, K-NN, and KD Tree algorithms, the satellite indices (i.e., DVI, NDVI, NDBI, UI and NBLI) were used to extract samples for model training and limit the need for ground truth data for validation of LULC maps of PEI. The results indicated that K-NN achieved an average kappa coefficient (82 and 74%) and high overall accuracy (86 and 81%) for sentinel-2A and Landsat-8, respectively. At the same time, the KD-Tree reached the average kappa coefficient (80 and 78%) and overall accuracy (85 and 84%) for Sentinel-2A and Landsat-8. Random forest achieved an average Kappa coefficient (83.3 and 73.3%) and high overall accuracy (92 and 85%) for Sentinel- 2A and Landsat-8 data, respectively, and proved a better classifier compared to K-NN and KD-Tree. The result of this objective also proved that vegetation indices have the potential to prepare the LULC maps.

The findings of Study 2 showed that employing multitemporal Planet imagery alone as an input to SVM and DT did not result in a meaningful difference in overall crop map accuracy. Multitemporal NDVI data combined with multitemporal Planet imagery, the overall accuracy of crop maps by SVM and DT enhanced. When integrated multitemporal Planet imagery with multitemporal NDVI data was used, the SVM algorithm-based prepared crop maps had a 6.25 % higher overall accuracy and a 7% higher kappa coefficient than DT.

The research 3 experimental activities were placed in a potato field on PEI (46.24136° N, 63.5311°W). For a real water need estimate, the potential of Planet imagery and GIS for creating new NDVI-based potato K_c, ETc maps at 3 m resolution was assessed. Between FAO suggested K_c and NDVI-based K_c values, a stronger correlation ($R^2 = 0.95$) was found. The produced potato ET_c maps revealed that ET_c was low at the start of growth, ranging from 0.69 to 2.01 mm/day, then increased to 1.89 to 3.05 mm/day as growth progressed. The largest ET_c value was found during the mid-stage of development, from 2.02 to 4.02 mm/day; subsequently, at the harvest stage of growth, from 1.06 to 2.16 mm/day, ET_c values began to decline. The developed potato ET_c maps to aid in estimating the variations in potato water requirements in a field across four development phases. The results presented above demonstrated that vegetation indices may be used to create LULC maps, improve crop mapping accuracy, and estimate potato ET_c.

REFERENCES

Abbas, Z., & Jaber, H. S. (2020). Accuracy assessment of supervised classification methods for extraction land use maps using remote sensing and GIS techniques. *IOP Conference Series: Materials Science and Engineering*, *745*(1), 0–13. https://doi.org/10.1088/1757-899X/745/1/012166

Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience and Remote Sensing*, *57*(1), 1–20. https://doi.org/10.1080/15481603.2019.1650447

Abubakar, G. A., Wang, K., Shahtahamssebi, A., Xue, X., Belete, M., Gudo, A. J. A., Shuka, K. A. M., & Gan, M. (2020). Mapping Maize Fields by Using Multi-Temporal Sentinel-1A and Sentinel-2A Images in Makarfi, Northern Nigeria, Africa. *Sustainability (Switzerland)*, *12*(6), 1–18. https://doi.org/10.3390/su12062539

Adamala, S., Rajwade, Y. A., & Reddy, Y. V. K. (2016). Estimation of wheat crop evapotranspiration using NDVI vegetation index. *Journal of Applied and Natural Science*, *8*(1), 159–166. https://doi.org/10.31018/jans.v8i1.767

Afzaal, H., Farooque, A. A., Abbas, F., Acharya, B., & Esau, T. (2020). Precision irrigation strategies for sustainable water budgeting of potato crop in prince Edward Island. *Sustainability (Switzerland)*, *12*(6). https://doi.org/10.3390/su12062419

Aghajanloo, M. B., Sabziparvar, A. A., & Hosseinzadeh Talaee, P. (2013). Artificial neural network-genetic algorithm for estimation of crop evapotranspiration in a semi-arid region of Iran. *Neural Computing and Applications*, *23*(5), 1387–1393. https://doi.org/10.1007/s00521-012-1087-y

Agriculture and land department of PEI. (2019). Accessed on 22 October 2020. https://www.princeedwardisland.ca/en/information/agriculture-and-land/agriculture-pei.

Agriculture and Agri-Food Canada Potato Market Information Review. (2016-2017). Accessed on 25 October 2020. https://www.agr.gc.ca/eng/horticulture/horticulture-sector-reports/potato-market-information-

Alface, A. B., Pereira, S. B., Filgueiras, R., & Cunha, F. F. (2019). Sugarcane spatial-temporal monitoring and crop coefficient estimation through NDVI | Monitoramento espaço-temporal da cana-de-açúcar e estimativa do coeficiente de cultivo por meio do NDVI. *Revista Brasileira de Engenharia Agricola e Ambiental*, *23*(5), 330–335.

Ali, M. Z., Qazi, W., & Aslam, N. (2018). A comparative study of ALOS-2 PALSAR and landsat-8 imagery for land cover classification using maximum likelihood classifier. *Egyptian Journal of Remote Sensing and Space Science*, *21*, 29–35. https://doi.org/10.1016/j.ejrs.2018.03.003

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotraspiration guidelines for computing crop water requirements. In *FAO Irrigation & drainage Paper 56. FAO, Food and Agriculture Organization of the United Nations,Roma*.

Allen, R. G., Pereira, L. S., Smith, M., Raes, D., & Wright, J. L. (2005). FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions. *Journal of Irrigation and Drainage Engineering*, *131*(1), 2–13. https://doi.org/10.1061/(ASCE)0733-9437(2005)131:1(2)

ALLEN2, P. D. & R. G. (2001). *Estimating reference evapotranspiration under inaccurate data conditions*. 33–45.

Asokan, A., Anitha, J., Ciobanu, M., Gabor, A., Naaji, A., & Hemanth, D. J. (2020). Image processing techniques for analysis of satellite images for historical maps classification-An overview. *Applied Sciences (Switzerland)*, *10*(12). https://doi.org/10.3390/app10124207

Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, *13*(1–2), 95–120. https://doi.org/10.1080/02757259509532298

Belova, L. (2018). *Could irrigation of potatoes grown for processing be economically feasible if high capacity wells are the source of water ?*

Boyle, S. A., Kennedy, C. M., Torres, J., Colman, K., Pérez-Estigarribia, P. E., & De La Sancha, N. U. (2014). High-resolution satellite imagery is an important yet underutilized resource in conservation biology. *PLoS ONE*, *9*(1), 1–11. https://doi.org/10.1371/journal.pone.0086908

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Burkhard, B., Kroll, F., Nedkov, S., & Müller, F. (2012). Mapping ecosystem service supply, demand and budgets. *Ecological Indicators*, *21*, 17–29. https://doi.org/10.1016/j.ecolind.2011.06.019

Cai, J., Liu, Y., Lei, T., & Pereira, L. S. (2007). Estimating reference evapotranspiration with the FAO Penman-Monteith equation using daily weather forecast messages. *Agricultural and Forest Meteorology*, *145*(1–2), 22–35. https://doi.org/10.1016/j.agrformet.2007.04.012

Chaves, M. E. D., Picoli, M. C. A., & Sanches, I. D. (2020). Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. *Remote Sensing*, *12*(18). https://doi.org/10.3390/rs12183062

Chen, D., Stow, D. A., & Gong, P. (2004). Examining the effect of spatial resolution and texture window size on classification accuracy: An urban environment case. *International Journal of Remote Sensing*, *25*(11), 2177–2192. https://doi.org/10.1080/01431160310001618464

Chen, S., Useya, J., & Mugiyo, H. (2020). Heliyon Decision-level fusion of Sentinel-1 SAR and Landsat 8 OLI texture features for crop discrimination and classi fi cation : case of Masvingo , Zimbabwe. *Heliyon*, *6*(August), e05358. https://doi.org/10.1016/j.heliyon.2020.e05358

Cheng, T., Ji, X., Yang, G., Zheng, H., Ma, J., Yao, X., Zhu, Y., & Cao, W. (2020). DESTIN: A new method for delineating the boundaries of crop fields by fusing spatial and temporal information from WorldView and Planet satellite imagery. *Computers and Electronics in*

Agriculture, *178*(September), 105787. https://doi.org/10.1016/j.compag.2020.105787

Cihlar, J. (2000). Land cover mapping of large areas from satellites: Status and research priorities. *International Journal of Remote Sensing*, *21*(6–7), 1093–1114. https://doi.org/10.1080/014311600210092

Clerici, N., Valbuena Calderón, C. A., & Posada, J. M. (2017). Fusion of sentinel-1a and sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia. *Journal of Maps*, *13*(2), 718–726. https://doi.org/10.1080/17445647.2017.1372316

Cover, T. M., & Hart, P. E. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964

de Oliveira, T. C., Ferreira, E., & Dantas, A. A. A. (2016). Variação temporal do índice de vegetação por diferença normalizada (NDVI) e obtenção do coeficiente de cultura (Kc) a partir do NDVI em áreas cultivadas com soja irrigada. *Ciencia Rural*, *46*(9), 1683–1688. https://doi.org/10.1590/0103-8478cr20150318

Department of Environment, E. and C. A. (n.d.). *Our Changing Climate*. Retrieved July 27, 2021, from https://www.princeedwardisland.ca/en/information/environment-energy-and-climate-action/our-changing-climate

Desai, C., & Umrikar, B. (2012). Image Classification Tool for Land Use Land Cover Analysis: A Comparative Study of Maximum Likelihood and Minimum Distance Method. 2 (3), 189-196.

Deschamps, B., McNairn, H., Shang, J., & Jiao, X. (2012). Towards operational radar-only crop type classification: Comparison of a traditional decision tree with a random forest classifier. *Canadian Journal of Remote Sensing*, *38*(1), 60–68. https://doi.org/10.5589/m12-012

Devadas, R., Denham, R. J., & Pringle, M. (2012). Support Vector Machine Classification of Object-Based Data for Crop Mapping, Using Multi-Temporal Landsat Imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information*

Sciences, *XXXIX-B7*(September), 185–190. https://doi.org/10.5194/isprsarchives-xxxix-b7-185-2012

Dolatshah, M., Hadian, A., & Minaei-Bidgoli, B. (2015). *Ball*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces*. *November*. http://arxiv.org/abs/1511.00628

Dombrowski, O., Franssen, H. H., & Brogi, C. (2021). *Performance of the ATMOS41 All-in-One Weather Station for Weather Monitoring*.

Doraiswamy, P. C., Hatfield, J. L., Jackson, T. J., Akhmedov, B., Prueger, J., & Stern, A. (2004). Crop condition and yield simulations using Landsat and MODIS. *Remote Sensing of Environment*, *92*(4), 548–559. https://doi.org/10.1016/j.rse.2004.05.017

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

El-Magd, I. A., & Tanton, T. W. (2003). Improvements in land use mapping for irrigated agriculture from satellite sensor data using a multi-stage maximum likelihood classification. *International Journal of Remote Sensing*, *24*(21), 4197–4206. https://doi.org/10.1080/0143116031000139791

Farg, E., Arafat, S. M., Abd El-Wahed, M. S., & El-Gindy, A. M. (2012). Estimation of Evapotranspiration ETc and Crop Coefficient K c of Wheat, in south Nile Delta of Egypt Using integrated FAO-56 approach and remote sensing data. *Egyptian Journal of Remote Sensing and Space Science*, *15*(1), 83–89. https://doi.org/10.1016/j.ejrs.2012.02.001

Fisette, T., Campbell, L., & Davidson, A. M. (2019). *The AAFC Annual Crop Inventory Context Agriculture in Canada and Globally*.

Fisette, T., Davidson, A., Daneshfar, B., Rollin, P., Aly, Z., & Campbell, L. (2014). Annual

space-based crop inventory for Canada: 2009-2014. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 5095–5098. https://doi.org/10.1109/IGARSS.2014.6947643

Fitting, T. P., & Models, Y. (2019). *Winter Wheat Yield Assessment from Landsat 8 and Sentinel-2 Data : Incorporating Surface Reflectance , Through Phenological Fitting , into Regression Yield Models*.

Fix, E. (1951). *Discriminatory analysis: nonparametric discrimination, consistency properties*.

Foody, G. M., Mathur, A., Sanchez-Hernandez, C., & Boyd, D. S. (2006). Training set size requirements for the classification of a specific class. *Remote Sensing of Environment*, *104*(1), 1–14. https://doi.org/10.1016/j.rse.2006.03.004

Forkuor, G., Conrad, C., Thiel, M., Ullmann, T., & Zoungrana, E. (2014). Integration of optical and synthetic aperture radar imagery for improving crop mapping in northwestern Benin, West Africa. *Remote Sensing*, *6*(7), 6472–6499. https://doi.org/10.3390/rs6076472

Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment*, *77*(3), 251–274. https://doi.org/10.1016/S0034-4257(01)00209-7

French, A. N., Hunsaker, D. J., Sanchez, C. A., Saber, M., Gonzalez, J. R., & Anderson, R. (2020). Satellite-based NDVI crop coefficients and evapotranspiration with eddy covariance validation for multiple durum wheat fields in the US Southwest. *Agricultural Water Management*, *239*(May), 106266. https://doi.org/10.1016/j.agwat.2020.106266

Glenn, E. P., Huete, A. R., Nagler, P. L., Hirschboeck, K. K., & Brown, P. (2007). Integrating remote sensing and ground methods to estimate evapotranspiration. *Critical Reviews in Plant Sciences*, *26*(3), 139–168. https://doi.org/10.1080/07352680701402503

Glenn, E. P., Neale, C. M. U., Hunsaker, D. J., & Nagler, P. L. (2011). Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems. *Hydrological Processes*, *25*(26), 4050–4062. https://doi.org/10.1002/hyp.8392

Gontia, N. K., & Tiwari, K. N. (2010). Estimation of crop coefficient and evapotranspiration of wheat (Triticum aestivum) in an irrigation command using remote sensing and GIS. *Water Resources Management*, *24*(7), 1399–1414. https://doi.org/10.1007/s11269-009-9505-3

Government of Prince Edward Island. (2020). *Prince Edward Island Population Report 2020*.

Hao, P., Zhan, Y., Wang, L., Niu, Z., & Shakir, M. (2015). Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. *Remote Sensing*, *7*(5), 5347–5369. https://doi.org/10.3390/rs70505347

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*. 562–578. https://doi.org/10.3390/rs2020562

Heydari, S. S., & Mountrakis, G. (2018). Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 Landsat sites. *Remote Sensing of Environment*, *204*(September), 648–658. https://doi.org/10.1016/j.rse.2017.09.035

Houborg, R., & McCabe, M. F. (2016). High-Resolution NDVI from planet's constellation of earth observing nano-satellites: A new data source for precision agriculture. *Remote Sensing*, *8*(9). https://doi.org/10.3390/rs8090768

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*(4), 725–749. https://doi.org/10.1080/01431160110040323

Illert, A., & Afflerbach, S. (2020). ISO 19131 AAFC Annual Crop Inventory - Data Product Specifications. In *Distribution* (Issue Iso 19131). http://www.agr.gc.ca/atlas/supportdocument_documentdesupport/aafcCropTypeMapping/en/ISO 19131_AAFC_Annual_Crop_Inventory_Data_Product_Specifications.pdf

Jamali, A. (2019). Evaluation and comparison of eight machine learning models in land use/land cover mapping using Landsat 8 OLI: a case study of the northern region of Iran. *SN Applied*

Sciences, *1*(11), 1–11. https://doi.org/10.1007/s42452-019-1527-8

Jia, K., Wei, X., Gu, X., Yao, Y., Xie, X., & Li, B. (2014). Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China. *Geocarto International*, *29*(8), 941–951. https://doi.org/10.1080/10106049.2014.894586

Jiang, H., Li, D., Jing, W., Xu, J., Huang, J., Yang, J., & Chen, S. (2019). Early season mapping of sugarcane by applying machine learning algorithms to sentinel-1A/2 time series data: A case study in Zhanjiang City, China. *Remote Sensing*, *11*(7). https://doi.org/10.3390/RS11070861

Karegowda, A. G., Jayaram, M. A., & Manjunath, A. S. (2012). Cascading K-means Clustering and K-Nearest Neighbor Classifier for Categorization of Diabetic Patients. *International Journal of Engineering and Advanced Technology*, *1*(3), 147–151.

Khosravi, I., Safari, A., Homayouni, S., & Mcnairn, H. (2017). Enhanced decision tree ensembles for land-cover mapping from fully polarimetric SAR data. *International Journal of Remote Sensing*, *38*(23), 7138–7160. https://doi.org/10.1080/01431161.2017.1372863

Kjaersgaard, J., Allen, R., & Irmak, A. (2011). Improved methods for estimating monthly and growing season ET using METRIC applied to moderate resolution satellite imagery. *Hydrological Processes*, *25*(26), 4028–4036. https://doi.org/10.1002/hyp.8394

Kumar, P., Gupta, D. K., Mishra, V. N., & Prasad, R. (2015). Comparison of support vector machine, artificial neural network, and spectral angle mapper algorithms for crop classification using LISS IV data. *International Journal of Remote Sensing*, *36*(6), 1604–1617. https://doi.org/10.1080/2150704X.2015.1019015

Kumar, P., Prasad, R., Choudhary, A., Mishra, V. N., Gupta, D. K., & Srivastava, P. K. (2017). A statistical significance of differences in classification accuracy of crop types using different classification algorithms. *Geocarto International*, *32*(2), 206–224. https://doi.org/10.1080/10106049.2015.1132483

Kussul, N., Lemoine, G., Gallego, F. J., Skakun, S. V., Lavreniuk, M., & Shelestov, A. Y.

(2016). Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(6), 2500–2508. https://doi.org/10.1109/JSTARS.2016.2560141

Kussul, N., Skakun, S., Shelestov, A., & Kussul, O. (2014). The use of satellite SAR imagery to crop classification in Ukraine within JECAM project. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *July*, 1497–1500. https://doi.org/10.1109/IGARSS.2014.6946721

Lei, H., & Yang, D. (2014). Combining the Crop Coefficient of Winter Wheat and Summer Maize with a Remotely Sensed Vegetation Index for Estimating Evapotranspiration in the North China Plain. *Journal of Hydrologic Engineering*, *19*(1), 243–251. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000765

Li, H., Wang, C., Zhong, C., Su, A., Xiong, C., Wang, J., & Liu, J. (2017). Mapping urban bare land automatically from Landsat imagery with a simple index. *Remote Sensing*, *9*(3). https://doi.org/10.3390/rs9030249

Li, H., Wang, C., Zhong, C., Zhang, Z., & Liu, Q. (2017). Mapping typical urban LULC from landsat imagery without training samples or self-defined parameters. *Remote Sensing*, *9*(7), 1–23. https://doi.org/10.3390/rs9070700

Li, P., & Moon, W. M. (2004). Land cover classification using MODIS–ASTER airborne simulator (MASTER) data and NDVI: A case study of the Kochang area, Korea. *Canadian Journal of Remote Sensing*, *30*(2), 123–136. https://doi.org/10.5589/m03-061

Li, W., Jiang, J., Guo, T., Zhou, M., Tang, Y., Wang, Y., Zhang, Y., Cheng, T., Zhu, Y., Cao, W., & Yao, X. (2019). Generating red-edge images at 3M spatial resolution by fusing sentinel-2 and planet satellite products. *Remote Sensing*, *11*(12), 1–18. https://doi.org/10.3390/rs11121422

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22.

Lin, B. B. (2010). The role of agroforestry in reducing water loss through soil evaporation and crop transpiration in coffee agroecosystems. *Agricultural and Forest Meteorology*, *150*(4), 510–518. https://doi.org/10.1016/j.agrformet.2009.11.010

Liu, J., Li, L., Huang, X., Liu, Y., & Li, T. (2019). *Mapping paddy rice in Jiangsu Province , China , based on phenological parameters and a decision tree model*. *13*(1), 111–123.

Lowe, B., & Kulkarni, A. (2015). Multispectral Image Analysis Using Random Forest. *International Journal on Soft Computing*, *6*(1), 1–14. https://doi.org/10.5121/ijsc.2015.6101

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870. https://doi.org/10.1080/01431160600746456

Mantero, P., Moser, G., & Serpico, S. B. (2004). Partially supervised classification of remote sensing images using SVM-based probability density estimation. *2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, *43*(3), 327–336. https://doi.org/10.1109/WARSD.2003.1295212

MacDonald, Mitch. (2020). A Billion Dollar Industry for PEI. Accessed 22 October 2020. https://www.theguardian.pe.ca/news/local/potatoes-a-billion-dollar-industry-for-pei.

Manuel, J. (2007). *Original article Discriminating cropping systems and agro-environmental measures*. *28*, 355–362.

Maponya, M. G., van Niekerk, A., & Mashimbye, Z. E. (2020). Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Computers and Electronics in Agriculture*, *169*(December 2019), 105164. https://doi.org/10.1016/j.compag.2019.105164

Maqsood, J., Farooque, A. A., Wang, X., Abbas, F., Acharya, B., & Afzaal, H. (2020). Contribution of climate extremes to variation in potato tuber yield in Prince Edward Island. *Sustainability (Switzerland)*, *12*(12). https://doi.org/10.3390/SU12124937

Matese, A., & Di Gennaro, S. F. (2021). Beyond the traditional NDVI index as a key factor to

mainstream the use of UAV in precision viticulture. *Scientific Reports*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-81652-3

Mather, P., & Tso, B. (2016). *Classification Methods for Remotely Sensed Data*. CRC Press. https://doi.org/10.1201/9781420090741

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. In *International Journal of Remote Sensing* (Vol. 39, Issue 9, pp. 2784–2817). Taylor and Francis Ltd. https://doi.org/10.1080/01431161.2018.1433343

McHugh, M. L. (2012). Lessons in biostatistics interrater reliability : the kappa statistic. *Biochemica Medica*, *22*(3), 276–282. https://hrcak.srce.hr/89395

Mohan, M., Pathan, S. K., Narendrareddy, K., Kandya, A., & Pandey, S. (2011). Dynamics of Urbanization and Its Impact on Land-Use/Land-Cover: A Case Study of Megacity Delhi. *Journal of Environmental Protection*, *02*(09), 1274–1283. https://doi.org/10.4236/jep.2011.29147

Morais, A., Fortin, V., & Anctil, F. (2015). Modelling of Seasonal Evapotranspiration from an Agricultural Field Using the Canadian Land Surface Scheme (CLASS) with a Pedotransfer Rule and Multicriteria Optimization. *Atmosphere - Ocean*, *53*(2), 161–175. https://doi.org/10.1080/07055900.2014.999745

Muavhi, N. (2020). *Evaluation of effectiveness of supervised classification algorithms in land cover classification using ASTER images-A case study from the Mankweng (Turfloop) Area and its environs , Limpopo Province , South Africa*. *9*(1).

Muraoka, H., Noda, H. M., Nagai, S., Motohka, T., Saitoh, T. M., Nasahara, K. N., & Saigusa, N. (2013). Spectral vegetation indices as the indicator of canopy photosynthetic productivity in a deciduous broadleaf forest. *Journal of Plant Ecology*, *6*(5), 393–407. https://doi.org/10.1093/jpe/rts037

Mustafa, S., Dauda, A. B., & Dauda, M. (2017). Image processing and SVM classification for

melanoma detection. *Proceedings of the IEEE International Conference on Computing, Networking and Informatics, ICCNI 2017*, *2017-Janua*, 1–5. https://doi.org/10.1109/ICCNI.2017.8123777

Narasimhulu, Y., Suthar, A., Pasunuri, R., & China Venkaiah, V. (2021). Ckd-tree: An improved kd-tree construction algorithm. *CEUR Workshop Proceedings*, *2786*, 211–218.

Nguyen, H. A. T., Sophea, T., Gheewala, S. H., Rattanakom, R., Areerob, T., & Prueksakorn, K. (2021). Integrating remote sensing and machine learning into environmental monitoring and assessment of land use change. *Sustainable Production and Consumption*, *27*, 1239–1254. https://doi.org/10.1016/j.spc.2021.02.025

Nguyen, H. T. T., Doan, T. M., & Radeloff, V. (2018). Applying Random Forest classification to map Land use/Land cover using Landsat 8 OLI. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(3W4), 363–367. https://doi.org/10.5194/isprs-archives-XLII-3-W4-363-2018

Nguyen, H. T. T., Doan, T. M., Tomppo, E., & McRoberts, R. E. (2020). Land use/land cover mapping using multitemporal sentinel-2 imagery and four classification methods-A case study from Dak Nong, Vietnam. *Remote Sensing*, *12*(9), 1–28. https://doi.org/10.3390/RS12091367

Nistor, M. M., Gualtieri, A. F., Cheval, S., Dezsi, Ş., & Boţan, V. E. (2016). Climate change effects on crop evapotranspiration in the Carpathian Region from 1961 to 2010. *Meteorological Applications*, *23*(3), 462–469. https://doi.org/10.1002/met.1570

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Onojeghuo, A. O., Blackburn, G. A., Wang, Q., Atkinson, P. M., Kindred, D., & Miao, Y. (2018). Mapping paddy rice fields by applying machine learning algorithms to multi-temporal sentinel-1A and landsat data. *International Journal of Remote Sensing*, *39*(4),

1042–1067. https://doi.org/10.1080/01431161.2017.1395969

Patel, J., Patel, H., & Bhatt, C. (2015). Modified Hargreaves Equation for Accurate Estimation of Evapotranspiration of Diverse Climate Locations in India. *Proceedings of the National Academy of Sciences India Section B - Biological Sciences*, *85*(1), 161–166. https://doi.org/10.1007/s40011-014-0314-y

Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., & Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, *115*(6), 1301–1316. https://doi.org/10.1016/j.rse.2011.01.009

Peña, J. M., Gutiérrez, P. A., Hervás-Martínez, C., Six, J., Plant, R. E., & López-Granados, F. (2014). Object-based image classification of summer crops with machine learning methods. *Remote Sensing*, *6*(6), 5019–5041. https://doi.org/10.3390/rs6065019

Phiri, D., & Morgenroth, J. (2017). Developments in Landsat land cover classification methods: A review. *Remote Sensing*, *9*(9). https://doi.org/10.3390/rs9090967

Piper, J. (1992). Variability and bias in experimentally measured classifier error rates. *Pattern Recognition Letters*, *13*(10), 685–692. https://doi.org/10.1016/0167-8655(92)90097-J

Planet Team. (2021). Planet Imagery Product Specifications. In *Planet Labs Inc*.

Rajah, P., Odindi, J., Mutanga, O., & Kiala, Z. (2019). The utility of Sentinel-2 Vegetation Indices (VIs) and Sentinel-1 Synthetic Aperture Radar (SAR) for invasive alien species detection and mapping. *Nature Conservation*, *35*, 41–61. https://doi.org/10.3897/natureconservation.35.29588

Ramezan, C. A., Warner, T. A., Maxwell, A. E., & Price, B. S. (2021). Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. *Remote Sensing*, *13*(3), 1–27. https://doi.org/10.3390/rs13030368

Rao, P., Zhou, W., Bhattarai, N., Srivastava, A. K., Singh, B., Poonia, S., Lobell, D. B., & Jain,

M. (2021). Using Sentinel-1, Sentinel-2, and Planet Imagery to Map Crop Type of Smallholder Farms. *Remote Sensing*, *13*(10), 1870. https://doi.org/10.3390/rs13101870

Reddy, C. (2012). Multi-Temporal Crop Classi cation Using a Decision Tree in a Southern Ontario Agricultural Region. In *Экономика Региона* (Issue July).

Reyes-Gonzalez, A., Hay, C., Kjaersgaard, J., & Neale, C. (2015). Use of remote sensing to generate crop coefficient and estimate actual crop evapotranspiration. *American Society of Agricultural and Biological Engineers Annual International Meeting 2015*, *6*(November 2016), 4374–4385. https://doi.org/10.13031/aim.20152190105

Reyes-González, A., Kjaersgaard, J., Trooien, T., Hay, C., & Ahiablame, L. (2018). Estimation of Crop Evapotranspiration Using Satellite Remote Sensing-Based Vegetation Index. *Advances in Meteorology*, *2018*(1). https://doi.org/10.1155/2018/4525021

Richardson, A. J., & Everitt, J. H. (1992). Using spectral vegetation indices to estimate rangeland productivity. *Geocarto International*, *7*(1), 63–69. https://doi.org/10.1080/10106049209354353

Richardson, A. J., & Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, *43*(12), 1541–1552.

Rodrigues, G. C., & Braga, R. P. (2021). A simple procedure to estimate reference evapotranspiration during the irrigation season in a hot-summer mediterranean climate. *Sustainability (Switzerland)*, *13*(1), 1–13. https://doi.org/10.3390/su13010349

Salovaara, K. J., Thessler, S., Malik, R. N., & Tuomisto, H. (2005). Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery. *Remote Sensing of Environment*, *97*(1), 39–51. https://doi.org/10.1016/j.rse.2005.04.013

Serra, P., More, G., & Pons, X. (2009). Thematic accuracy consequences in cadastre land-cover enrichment from a pixel and from a polygon perspective. *Photogrammetric Engineering and Remote Sensing*, *75*(12), 1441–1449. https://doi.org/10.14358/PERS.75.12.1441

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017). Exploring Google earth engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science*, *5*, 1–10. https://doi.org/10.3389/feart.2017.00017

Shivakumar, B. R., & Rajashekararadhya, S. V. (2018). Investigation on land cover mapping capability of maximum likelihood classifier: A case study on North Canara, India. *Procedia Computer Science*, *143*, 579–586. https://doi.org/10.1016/j.procs.2018.10.434

Shock, C. C., Feibert, E. B. G., & Saunders, L. D. (1998). Potato yield and quality response to deficit irrigation. In *HortScience* (Vol. 33, Issue 4, pp. 655–659). https://doi.org/10.21273/hortsci.33.4.655

Showqi, I., Rashid, I., & Romshoo, S. A. (2014). Land use land cover dynamics as a function of changing demography and hydrology. *GeoJournal*, *79*(3), 297–307. https://doi.org/10.1007/s10708-013-9494-x

Soil organic matter status on PEI. (2018). Accessed 4 November 2020. https://www.princeedwardisland.ca/en/information/agriculture-and-land/soil-organic-matter-status-pei

Simonneaux, V., Duchemin, B., Helson, D., Er-Raki, S., Olioso, A., & Chehbouni, A. G. (2008). The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *International Journal of Remote Sensing*, *29*(1), 95–116. https://doi.org/10.1080/01431160701250390

Sitokonstantinou, V., Papoutsis, I., & Kontoes, C. (2018). *Scalable Parcel-Based Crop Identification Scheme Using Sentinel-2 Data Time-Series for the Monitoring of the Common Agricultural Policy*. *Cc*. https://doi.org/10.3390/rs10060911

Story, M., & Congalton, R. G. (1986). Remote Sensing Brief Accuracy Assessment: A User's Perspective. *Photogrammetric Engineering and Remote Sensing*, *52*(3), 397–399.

Tasumi, M., & Allen, R. G. (2007). Satellite-based ET mapping to assess variation in ET with

timing of crop development. *Agricultural Water Management*, *88*(1–3), 54–62. https://doi.org/10.1016/j.agwat.2006.08.010

Thanh Noi, P., & Kappas, M. (2017). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors (Basel, Switzerland)*, *18*(1). https://doi.org/10.3390/s18010018

Topaloğlu, R. H., Sertel, E., & Musaoğlu, N. (2016). Assessment of classification accuracies of Sentinel-2 and Landsat-8 data for land cover/use mapping. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *41*(5), 1055–1059. https://doi.org/10.5194/isprsarchives-XLI-B8-1055-2016

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

United Nations Department of Economic and Social Affairs_Population Division. (2021). Global Population Growth and Sustainable Development. In *Un Desa/Pop/2021/Tr/No. 2*. www.unpopulation.org.

Vajda, S., & Santosh, K. C. (2017). A fast k-nearest neighbor classifier using unsupervised clustering. *Communications in Computer and Information Science*, *709*, 185–193. https://doi.org/10.1007/978-981-10-4859-3_17

Valipour, M., & Eslamian, S. (2014). Analysis of potential evapotranspiration using 11 modified temperature-based models. *International Journal of Hydrology Science and Technology*, *4*(3), 192–207. https://doi.org/10.1504/IJHST.2014.067733

Verma, A. K. (2017). Sugarcane crop identification from LISS IV data using ISODATA , MLC , and indices based decision tree approach. *Arabian Journal of Geosciences*. https://doi.org/10.1007/s12517-016-2815-x

Vibhute, A. D., & Gawali, B. W. (2013). Analysis and Modeling of Agricultural Land use using Remote Sensing and Geographic Information System: a Review. *International Journal of*

Engineering Research and Applications, *3*(3), 81–91.

Vishwanath, N., Ramesh, B., & Sreenivasa Rao, P. (2016). Unsupervised classification of remote sensing images using k-means algorithm. *INTERNATIONAL JOURNAL OF LATEST TRENDS IN ENGINEERING AND TECHNOLOGY*, *7*(2). https://doi.org/10.21172/1.72.584

Vladimirova, T., Wu, X., & Bridges, C. P. (2008). Development of a satellite sensor network for future space missions. *IEEE Aerospace Conference Proceedings*. https://doi.org/10.1109/AERO.2008.4526248

Wang, C., Du, J., Chen, G., Wang, H., Sun, L., Xu, K., Liu, B., & He, Z. (2019). QAM classification methods by SVM machine learning for improved optical interconnection. *Optics Communications*, *444*(January), 1–8. https://doi.org/10.1016/j.optcom.2019.03.058

Watts, J. D., & Lawrence, L. R. (2008). Merging Random Forest Classification With an Object-Oriented Approach for Analysis of Agricultural Lands. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *37*, 2018.

Xiang, K., Li, Y., Horton, R., & Feng, H. (2020). Similarity and difference of potential evapotranspiration and reference crop evapotranspiration – a review. *Agricultural Water Management*, *232*(January). https://doi.org/10.1016/j.agwat.2020.106043

Xiao, X., Boles, S., Liu, J., Zhuang, D., & Liu, M. (2002). *Characterization of forest types in Northeastern China , using multi-temporal SPOT-4 VEGETATION sensor data*. *82*, 335–348.

Yang, N., Liu, D., Feng, Q., Xiong, Q., Zhang, L., Ren, T., Zhao, Y., Zhu, D., & Huang, J. (2019). Large-scale crop mapping based on machine learning and parallel computation with grids. *Remote Sensing*, *11*(12), 1–22. https://doi.org/10.3390/rs11121500

Yi, Z., Jia, L., & Chen, Q. (2020). Crop classification using multi-temporal sentinel-2 data in the Shiyang river basin of China. *Remote Sensing*, *12*(24), 1–21. https://doi.org/10.3390/rs12244052

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*(3), 583–594. https://doi.org/10.1080/01431160304987

ZHANG, H., XU, F., WU, Y., HU, H. hai, & DAI, X. feng. (2017). Progress of potato staple food research and industry development in China. *Journal of Integrative Agriculture*, *16*(12), 2924–2932. https://doi.org/10.1016/S2095-3119(17)61736-2

Zhang, Y., Han, W., Niu, X., & Li, G. (2019). Maize crop coefficient estimated from UAV-measured multispectral vegetation indices. *Sensors (Switzerland)*, *19*(23), 1–17. https://doi.org/10.3390/s19235250

Zhu, L., Radeloff, V. C., & Ives, A. R. (2017). Improving the mapping of crop types in the Midwestern U.S. by fusing Landsat and MODIS satellite data. *International Journal of Applied Earth Observation and Geoinformation*, *58*, 1–11. https://doi.org/10.1016/j.jag.2017.01.012