ERDC/GRL TR-21-7

Geospatial Research Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

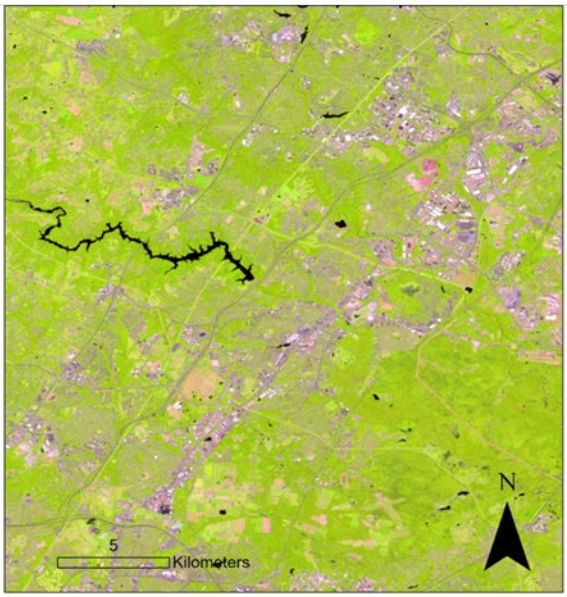
ERDC
ENGINEER RESEARCH & DEVELOPMENT CENTER

Semi-Automated Land Cover Mapping Using an Ensemble of Support Vector Machines with Moderate Resolution Imagery Integrated into a Custom Decision Support Tool

Dr. Kristofer Lasko, Dr. Elena Sava

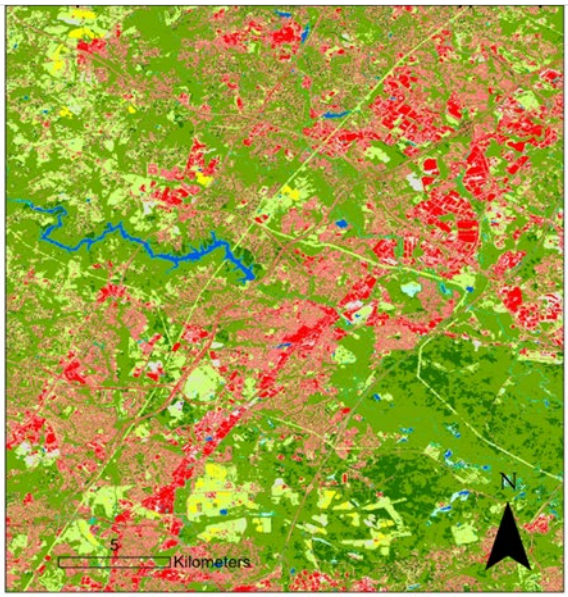
November 2021

Input: Satellite Imagery Composite



Sentinel-2 Imagery (20m), Laurel, MD, July 2019
R: SWIR-1
G: Near-IR
B: Red

Output: Semi-automated Land Cover Map



- Unclassified
- Water
- Bare Ground
- Low Vegetation
- Evergreen Trees
- Deciduous Trees
- Herbaceous Wetlands
- Woody Wetlands
- Temporary Water
- Built-up
- Built-up (Low Intensity)

The U.S. Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdc.usace.army.mil.

To search for other technical reports published by ERDC, visit the ERDC online library at https://erdclibrary.on.worldcat.org/discovery.

ERDC/GRL TR-21-7
November 2021

Semi-Automated Land Cover Mapping Using an Ensemble of Support Vector Machines with Moderate Resolution Imagery Integrated into a Custom Decision Support Tool

Dr. Kristofer Lasko, Dr. Elena Sava

Geospatial Research Laboratory
U.S. Army Engineer Research and Development Center
7701 Telegraph Road
Alexandria, VA 22315-3864

Final Report

Prepared for Geospatial Research Laboratory
U.S. Army Engineer Research and Development Center
7701 Telegraph Road, Alexandria, VA 22315-3864

Under PE 633463/Project AU1/Tactical Geospatial Information Capabilities,
Enhanced Terrain Processing—Demonstration

Abstract

Land cover type is a fundamental remote sensing-derived variable for terrain analysis and environmental mapping applications. The currently available products are produced only for a single season or a specific year. Some of these products have a coarse resolution and quickly become outdated, as land cover type can undergo significant change over a short time period. In order to enable on-demand generation of timely and accurate land cover type products, we developed a sensor-agnostic framework leveraging pre-trained machine learning models. We also generated land cover models for Sentinel-2 (20m) and Landsat 8 imagery (30m) using either a single date of imagery or two dates of imagery for mapping land cover type. The two-date model includes 11 land cover type classes, whereas the single-date model contains 6 classes. The models' overall accuracies were 84% (Sentinel-2 single date), 82% (Sentinel-2 two date), and 86% (Landsat 8 two date) across the continental United States. The three different models were built into an ArcGIS Pro Python toolbox to enable a semi-automated workflow for end users to generate their own land cover type maps on demand. The toolboxes were built using parallel processing and image-splitting techniques to enable faster computation and for use on less-powerful machines.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the Geospatial Research Laboratory under PE 633463, Project AU1, Tactical Geospatial Information Capabilities, Enhanced Terrain Processing—Demonstration.

The work was performed under the Data and Signature Analysis Branch (TRS) of the TIG Research Division (TR), U.S. Army Engineer Research and Development Center, Geospatial Research Laboratory (ERDC-GRL). At the time of publication, Ms. Jennifer Smith was Chief, Data Signature and Analysis Branch (CEERD-TRS); Mr. Jeffrey Murphy was Chief, TIG Research Division (CEERD-TR); and Mr. Vijay Acharya was Director of the Technical Directorate (CEERD-TET). The Deputy Director of ERDC-GRL was Ms. Valerie L. Carney, and the Director was Mr. Gary Blohm.

COL Teresa Schlosser was Commander of ERDC, and Dr. David W. Pittman was the Director.

Acronyms and Abbreviations

Acronym	Meaning
ARD	Analysis Ready Data
CONUS	Continental United States
CPU	Central processing unit
DoD	Department of Defense
ETP	Enhanced terrain processing
GDAL	Geospatial Data Abstraction Library
GRL	Geospatial Research Laboratory
GLC	Global land cover
GUI	Graphical user interface
HPC	High performance computing
OCONUS	Outside Continental United States
L2A	Level-2A (surface reflectance)
NIR	Near-infrared
NDVI	Normalized-Difference Vegetation Index
NLCD	National Land Cover Database
OLI	Operational Land Imager
RGB	Red, green, blue
SWIR	Shortwave infrared
SVM	Support vector machine
TIRS	Thermal Infrared Sensor
USGS	United States Geological Survey
2D	Two-dimensional
3D	Three-dimensional

1 Introduction

1.1 Background

The spatial patterns of land cover reflect the underlying natural and social processes of an area and thus provide essential information for many applications and disciplines. Land cover type is a fundamental variable for characterizing ecosystems, understanding demographics, performing terrain analysis, responding to natural hazards, and much more. With the ability to provide precise and recurrent observations, remote sensing imagery has long been key to monitoring land cover. Remotely sensed imagery from publicly available satellites such as Sentinel-2 and Landsat 8 offers consistent observations in both near-real time and historical imagery. These instruments offer combined repeat land observations several times per month, enabling the characterization of land cover and associated changes across multiple time periods. Consistently available satellite imagery at moderate resolution enables timely monitoring of land cover and associated change. Monitoring of land cover is of critical importance in both civil and military applications.

Typically, land cover maps are created using supervised classification methods where the user manually labels pixels of representative land cover features in imagery and then uses the imagery and associated labels to train a classifier in order to produce a thematic map. Land cover type maps are not new and have been produced at local, regional, and global scales for many years. For example, at the local scale, land cover maps often include details about crop-specific land cover types such as wheat and corn and in Kyiv, Ukraine (Kussul et al. 2017), triple-, double-, and single-cropped rice in the Mekong Delta, Vietnam (Kontgis, Schneider, and Ozdogan 2015), or in the Red River Delta, Vietnam (Lasko et al. 2018). At the regional and global scales, land cover maps represent multiple land cover types as the spectral signatures of each land cover vary significantly across environmental conditions. Such maps often provide broad stratification of diverse spatial patterns. For example, at the 30m Landsat scale, regional cropland extent maps have been produced for Australia, China, and continental Africa (Xiong et al. 2017; Teluguntla et al. 2018). These maps are binary in nature and indicative of areas that are either cropland or not cropland and thus provide more generalized categories. In order to provide timely and accurate land cover products at a global scale,

hundreds of thousands of satellite images must be analyzed, and High Performance Computing (HPC) such as Google Earth Engine must be leveraged to effectively process large quantities of data. For example, both Hansen et al. (2013) and Pickens et al. (2020) leverage HPC approaches to create 30m land cover maps highlighting temporal changes in forest cover (Hansen et al.) and surface water seasonality and extent (Pickens et al.) at a global scale. Oftentimes, global and regional land cover products have significant barriers to overcome in order to provide products that are locally relevant or accurate. These particular studies aimed to overcome this obstacle through moderate resolution time series of imagery and robust machine learning classification methods.

A selection of regional and global land cover maps exist based on moderate resolution 30m Landsat imagery. One of the most popular and robust products is the National Land Cover Database (NLCD) produced at intermittent annual time periods using a full time series of imagery (Jin et al. 2019). It is currently available for 2001, 2003, 2006, 2008, 2011, 2013, and 2016. The 2016 NLCD provides the following land cover categories for the continental United States (CONUS): open water, perennial snow/ice, developed open space, developed low intensity, developed medium intensity, developed high intensity, barren land, deciduous forest, evergreen forest, mixed forest, shrub forest, herbaceous forest, shrub, grassland/herbaceous, pasture hay, cultivated crops, woody wetlands, and herbaceous wetlands. Another regional general land cover product is the Corine Land Cover database at 100m resolution produced across Europe with high fidelity and accuracy (Buttner 2014). Ten class global land cover (GLC) maps for the years 2000 and 2010 were produced at 30m scale based on thousands of Landsat and HuanJing-1 satellite images.

Land cover products from such robust datasets are used as input in a variety of applications and analyses ranging from hydrology and watershed management to agricultural studies as well as urban and social studies. In order to produce robust products such as NLCD or the Corine Land Cover database, a dense time series stack (15-plus dates of imagery) and advanced sampling methods are required. This requires significant computing capability, storage space, and time. More recently, research has leveraged machine learning and existing training datasets to create regional land cover maps with high accuracy and minimal time spent generating training data (Malinowski et al. 2020).

While there have been many studies conducted at high spatial resolution, most of the available regional and global land cover products are at 30m spatial resolution or coarser. Some are available at higher resolution but often have limitations such as the need for dozens of time series images in order to create the land cover product (Malinowski et al. 2020). Reliance on static maps at time scales that quickly become out of date (e.g., the latest NLCD map was produced in 2016 but the current year is 2021) means that land surface monitoring applications may be inaccurate in areas experiencing land cover change on an annual basis. To resolve this issue, decision support tools could be created to enable the user to create their own maps during their selected time period and with imagery they can independently acquire. This would cut down on dataset size and would save analysts significant processing time.

This technical report describes a methodology using freely accessible satellite imagery and pregenerated training labels to create land cover classification models using machine learning. The models are then wrapped into a Python script and directly integrated into ArcGIS Pro as a Python toolbox. These models enable automatic classification by the end user without the need to input any manually created training labels. Three different land cover mapping models and tools were created: (1) single-date Sentinel-2 classifier, (2) two-date Sentinel-2 classifier, and (3) two-date Landsat 8 classifier. These tools are a part of the Enhanced Terrain Processing Toolkit (ETPT), which is a custom toolbox within ESRI ArcGIS Pro designed for processing high-resolution data and generating timely geospatial products to support actionable maneuver and tactical force protections. As input, the semiautomatic land cover tools can exploit the latest available Sentinel-2 or Landsat 8 multispectral imagery to generate timely geospatial products at moderate spatial resolution and improved temporal resolutions, enabling the end user to have a timely and accurate land cover product.

1.2 Image classification

Thematic mapping from remotely sensed imagery is typically based on two main classification methods: unsupervised and supervised. Unsupervised classification seeks to group together classes by their relative spectral similarity, and a user must interpret how the clusters are related. Supervised classification aims to classify pixels on the basis of their spectral similarity to a set of meaningful training labels provided by a human. Supervised classification is more time consuming due to the

creation of feature labels, but often results in more meaningful output and is the prevalent method today for land cover mapping as it often has higher accuracy (Hasmadi, Pakhriazad, and Shahrin 2009). A variety of supervised classification algorithms is used in land cover mapping studies. These include support vector machines (SVM), ensembles of weak learner decision tree methods such as random forest, various types of artificial neural networks, probabilistic classifiers such as Naïve Bayes, and much more with linear and nonlinear models (Pedregosa et al. 2011). In this study, we implemented an ensemble of three linear SVM classifiers.

1.2.1 Support vector machines

The SVM classifier is a widely used nonparametric statistical learning classifier with no assumptions made regarding the underlying data distribution. SVMs generally perform well on sparse data, and they are also less likely to overfit the model than a decision tree classifier such as random forest. The SVM is versatile in that different kernel functions can be used for training the model, such as linear, sigmoid, or polynomial. While the SVM was originally designed for binary classification problems, the SVM has been shown to be effective in high dimensional spaces. This method typically performs effectively in land cover classification studies using shallow learning techniques (Foody and Mathur 2004; Pal and Mather 2005; Jia et al. 2012; Pal and Foody 2012). The SVM algorithm promises to obtain the optimal hyperplane for a training dataset in terms of the generalization error. A detailed description of the SVM algorithm can be found in Suykens and Vandewalle (1999). Additional description of our classification method is in the methodology section.

1.2.2 Classifier training

Classifier training starts with the process of collecting satellite imagery and labeling pixels with land cover classes of interest. This is typically a manually intensive procedure that involves many hours of work by an analyst. One of the most accurate methods of training data selection is to extract data proportional to the expected occurrence of each respective land cover class (Zhu et al. 2016). However, it is also critical that upper and lower bounds be set for the amount of training data to select per class so that a classification model is not overwhelmed and so that enough representative data is extracted from low proportion classes (Olofsson et al. 2014; Zhu et al. 2016; Congalton and Green 2019).

Training data collection can be done following stratified random sampling approaches or it can be collected strategically by expert image interpretation. The advantage of the former is that a portion of the training data can be withheld from training the classifier. This portion can then be used to evaluate the accuracy of the classifier in a robust manner without the need to generate additional points for accuracy assessment. Strategic generation of training data is also effective if done methodically so as not to over fit the model but requires additional work to create a robust accuracy assessment. Strategic collection is also useful when the land cover proportions are not well known a priori.

Training data is typically generated through a labor-intensive process where an analyst uses multiple sources to make an informed decision about the land cover label. It is critical that the analyst generate the training points or polygons directly on the imagery to be used for classifier training. If labels are created on the supplemental imagery, it can result in geolocation errors that affect the labeling and associated pixel values (Congalton and Green 2019). Supplemental imagery sources that can assist training data collection include higher-resolution imagery, time series of imagery with data collected at key phenology dates, Google Earth base maps and field- or ground-collected geotagged photos from social media or an institution. Imagery can also be supplemented by other authoritative sources such as other thematic maps of the regions, peer-reviewed articles, and other reputable nonimagery sources that can convey information about the region of interest.

Classifier model training occurs after a sufficient amount of training data is collected, typically not more than 15,000 to 20,000 pixels for a given Landsat-sized scene. During classifier training, the pixel values for each spectral band along with the associated land cover class labels (e.g., evergreen trees) are combined into a single tabular file, verified, and passed into a machine learning classifier such as an SVM. Oftentimes, it is necessary to collect additional training data as initial classifier training often reveals errors.

1.2.3 Classification accuracy assessment

An initial assessment of accuracy can be gleaned through withholding a portion of the training data to be used for accuracy assessment. This data is used to test the trained classifier and compare the model predicted output to our ground-truth labels. From this test data, we can quickly

compute preliminary accuracy metrics based on the resulting error matrix, which shows the number of correctly classified pixels for each class. The most useful metrics at this stage are class-specific omission errors, class-specific commission errors, and overall accuracy. Additional training data can be generated in land cover categories with relatively high omission or commission errors. The error matrix can also show which classes are often confused, identify which can also be used to create additional training data, or revisit old data for verification. While it is often recommended to aim for an overall accuracy above 80%, this standard is arbitrary and dependent upon the difficulty of the classification task at hand. Higher accuracies should be expected for simple classification tasks such as vegetation versus nonvegetation, whereas lower accuracies may be expected for complex classification schemes with multiple classes or difficult-to-differentiate classes across a large area of interest. For example, the NLCD 2016 product with Anderson level II classes had an estimated overall accuracy of 82% from 4 selected scenes (Jin et al. 2019). Whereas for a smaller area with only 7 classes, higher accuracy can be achieved, such as in the case of mapping in Ukraine with overall accuracy in excess of 90% (Lavreniuk et al. 2015).

1.3 Objectives

Creation of image labels and associated training data for generating land cover type maps is a time consuming and iterative process. When training data is created as a part of disparate projects, it results in data inconsistencies that propagate into the final land cover type map. Moreover, reliance on coarse resolution (30m or coarser), static date land cover maps is costly and delivers a product that quickly becomes outdated and obsolete to the present time. Based on this, our objectives are to (1) generate land cover type training data across the CONUS, (2) build a supervised classification model to enable on-demand land cover generation across the United States without a need for the user to generate their own data, (3) test and evaluate the accuracy of the model, and (4) combine the model and workflow into a Python-based toolbox and integrate it into ArcGIS Pro software to enable endusers to automatically generate their own land cover maps without the need for their own training data. The ultimate goal of this project, beyond this specific study, is to build a sensor-agnostic land cover classification model and associated decision support tool. This report will discuss the overall sensor-agnostic methodology that will be implemented as part of a long-term objective, but

the focus will be on the aforementioned objectives relating to the work that has been completed.

2 Data and Training Sites

2.1 Landsat 8 and Sentinel-2 imagery

The ability to have continuous satellite data available at a large spatial extent requires a constellation of satellites that have long-term data records as well as new satellite constellations that collect data at frequent intervals. For these reasons, as well as global coverage and free data access, Landsat 8 and Sentinel-2 were selected for this study.

Landsat is NASA's longest-enduring Earth-observing mission and contains historical imagery dating back to the early 1970s from several different missions. The latest operational Landsat 8 satellite launched in 2013 has a 16-day temporal resolution and consists of two instruments on board, Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The two active sensors together provide seasonal coverage of the global landmass at various spatial resolutions ranging from 30m (visible, near-infrared [NIR], shortwave infrared [SWIR]), to 100m in the TIRS1 and TIRS2 and 15m (panchromatic). The OLI instrument monitors the Earth's surface in the visible, NIR, and SWIR sections of the electromagnetic spectrum and is commonly used for land cover or vegetation change mapping. A series of Landsat 8 Analysis Ready Data (ARD) surface reflectance scenes were used based on data availability for our analysis and cloud presence in each image.

The European Space Agency and the European Commission started a constellation initiative with the launch of the first Sentinel-2 mission in April 2014. Together with other satellite platforms that have a wider spatial coverage such as Landsat 8, the number of observations for an area increased. Currently, two operational Sentinel-2 satellites are available with multispectral instruments to acquire data at a moderate temporal and moderate spatial resolution. Sentinel-2 collects data at a 10m, 20m, and 60m spatial resolution depending on the spectral band. It has a temporal resolution of approximately 5 days with two satellites in constellation. Sentinel-2 imagery can be acquired from a variety of US sources or directly from the European Space Agency’s public-facing servers available at the open access hub website. Each scene file contains 13 multispectral spectral bands ranging from the visible, NIR, and SWIR electromagnetic frequency domain. Sentinel-2 land surface reflectance scenes at Level 2A (L2A) were downloaded and used for the land cover training data collection and model

development. Table 1 summarizes the spatial and spectral characteristics of Landsat 8 ARD and Sentinel-2 L2A datasets used to develop the land cover classification tool in this study.

Table 1. Multispectral characteristics of Landsat 8 (ARD surface reflectance) and Sentinel-2 (L2A surface reflectance) satellite spectral bands used in this study.

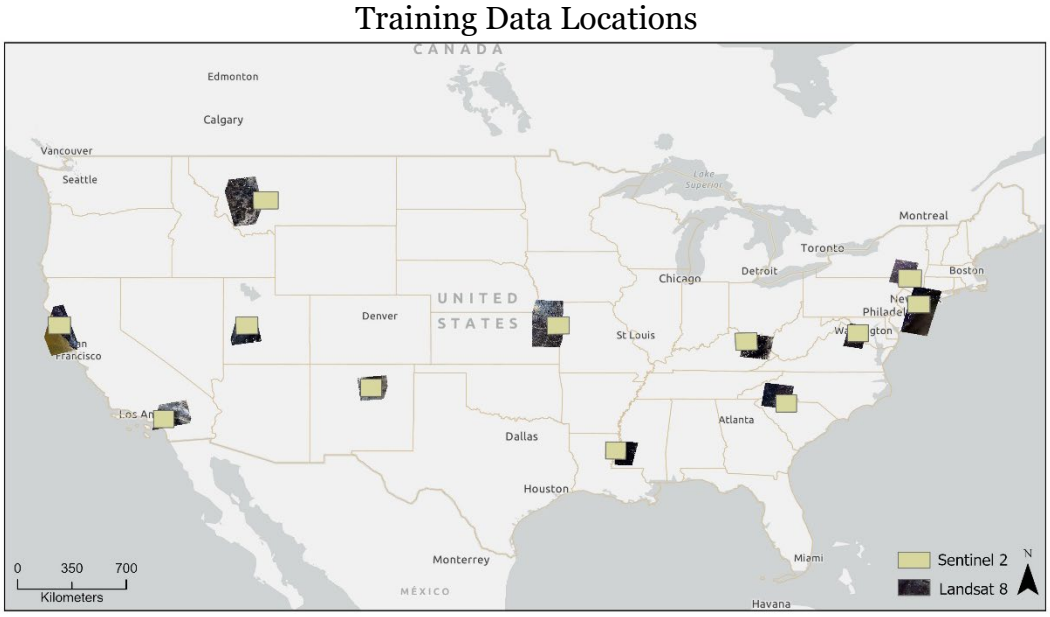
Landsat 8 OLI		Sentinel-2 MSI	
Bands Spectral Region	Resolution (m)	Bands Spectral Region	Resolution(m)
Band 2—Blue	30	Band 2—Blue	20
Band 3—Green	30	Band 3—Green	20
Band 4—Red	30	Band 4—Red	20
Band 5—NIR	30	Band 5—Red Edge	20
Band 6—SWIR1	30	Band 6—Red Edge	20
Band 7—SWIR2	30	Band 7—Red Edge	20
-	-	Band 8A—NIR Narrow	20
-	-	Band 11—SWIR	20
-	-	Band 12—SWIR	20

2.2 Description of training sites

Landsat 8 and Sentinel-2 multispectral imagery were collected throughout different sites within the CONUS to represent the diverse land cover types and ecosystems as guided by the 2016 NLCD product with training locations shown in Figure 1. The reasons for selecting these sites follows in the proceeding sentences. The northern California site is representative of lush evergreen forests found throughout the west coast and secondarily includes a mosaic of low vegetation, water, and small suburban areas. The southern California site includes the dry urban area of Los Angeles as well as arid landscape with less dense forests, bare areas, and croplands. The Utah site is characteristic of arid, rocky, salt-laden landscape but also includes significant areas of cropland and built-up areas. The Mississippi and Louisiana site is characteristic of humid subtropical forest, wetlands, irrigated croplands, lush mixed forest, and the Mississippi Valley landscape. The New York site is representative of temperate continental climate landscape with a mosaic of low vegetation, deciduous and evergreen forest, lakes, and more. The New Jersey site is representative of warm continental climate with wetlands, built-up areas, forests, and more. The Virginia, Maryland, DC scene is representative of warm oceanic

climate with a mosaic of densely built-up areas, suburbs, wetlands, rivers, croplands, deciduous forest, and more. The Kentucky and South Carolina scenes are representative of mosaic landscapes in cool continental and humid subtropical environments, respectively. The Kansas scene is representative of large-scale commercial cropland such as soybean and corn. The Western Montana site is a mosaic of plains, forest, wetlands, and croplands. Lastly, the Arizona site is characteristic of semi-arid higher elevation areas. Figure 1 shows sites where Landsat 8 and Sentinel-2 images were collected.

Figure 1. Map showing the locations of Sentinel-2 and Landsat 8 training data sites used for building the semi-automated land cover classification models.



2.3 Land cover datasets

Several existing land cover products were acquired for comparison and evaluation purposes with respect to the maps produced in this study. These products have been ground-truthed and have been used in many different studies.

The Geocover land cover product was produced using Landsat imagery from the 1990s and early 2000s based on the expert assignment of clusters resulting from unsupervised classification (Cunningham et al. 2002). The NLCD product was produced for year 2016 using supervised machine learning classification from a dense time series of Landsat satellite observations. This product is freely available from the United States

Geological Survey (USGS) Earth Explorer website (https://earthexplorer.usgs.gov/). For this study, we acquired both the NLCD and Geocover datasets over our test sites for comparison with our results.

3 Methodology

3.1 Methodology

The goal of this project is to create a suite of decision support tools for automated land cover type mapping. The underlying framework of the tools rely on pretrained machine learning models that enable the tools to be used without the need for additional training data or tedious user input or training data collection. Such tools allow the end user to generate up-to-date land cover type maps at higher temporal and spatial resolutions in comparison to outdated and coarser land cover products. Figure 2 highlights the methodology used to achieve this goal. Sentinel-2 and Landsat 8 imagery are used to create training data for selected land cover types across different ecological regions within CONUS. This data is then used as input to train several machine learning classifiers, specific for Sentinel-2 and another for Landsat 8. After each classifier is built, the decision from three SVMs is combined to improve overall performance, forming an ensemble of SVM machine learning classifiers. These classification models are wrapped into a Python toolbox and enable the end user to leverage them for automated image classification in their area of interest. In this technical report, the focus is on generating rapid land cover type mapping models and decision support tools for Landsat 8 and Sentinel-2 satellite imagery, as well as the potential development for a sensor-agnostic model.

Figure 2. Depiction of underlying methodology for each land cover classification.

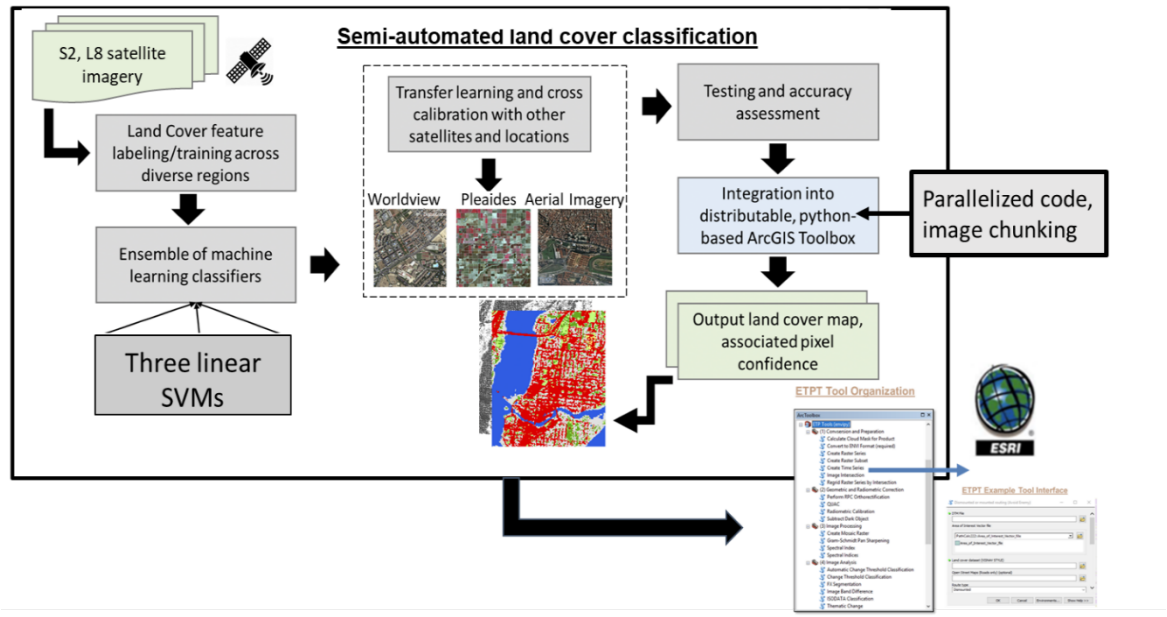
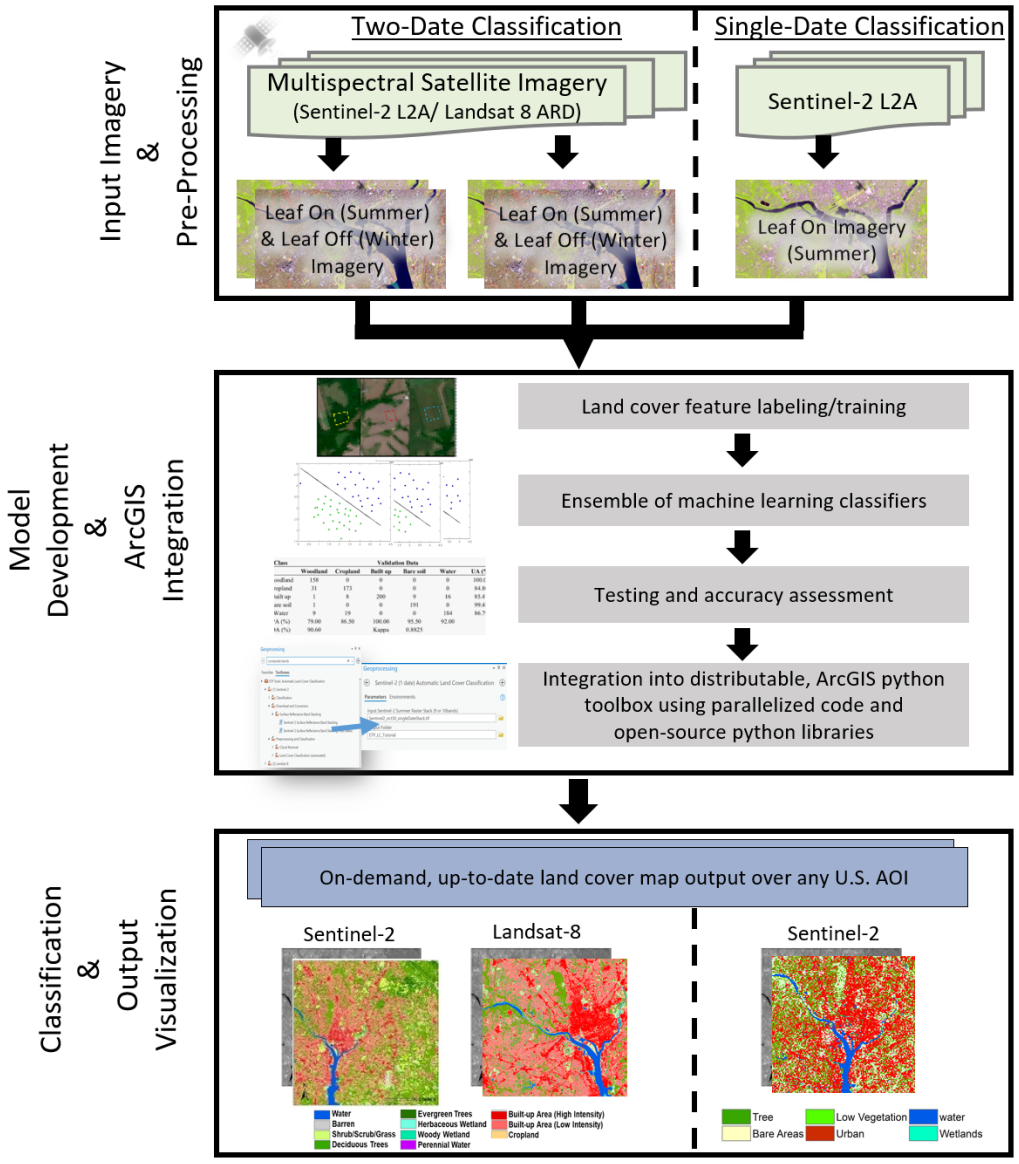


Figure 3 highlights the overview of the methodology. As mentioned in the data section, we acquired Sentinel-2 and Landsat 8 images across twelve diverse sites of the CONUS representative of the varying terrain, climates, and resulting variation in land cover.

Figure 3. Overview of the methodology developed for this project which leverages Sentinel-2 and Landsat 8 imagery to create land cover type maps.



Training polygons for both Landsat 8 and Sentinel-2 were carefully and strategically created across the land cover types of interest. This data was used by the classification algorithm to recognize and understand patterns and create information that it will use for predicting land cover type. Separate training polygons were created for each of the machine learning models used to build the semiautomatic land cover tools, single-date Sentinel-2 classifier, two-date Sentinel-2 classifier, and two-date Landsat 8 classifier. The single-date classifier uses imagery from the spring/summer season, whereas the two-date classifiers use an image

from Spring/Summer (leaf-on) and one from Fall/Winter (leaf-off). The additional imagery enables more land cover classes.

With regard to the training data creation for the two-date classifiers, both dates of imagery along with high-resolution basemaps and Google Earth Pro historical imagery were together used to make an expert assessment of the land cover type for each polygon. Hundreds of small polygons were created across each land cover type as opposed to several large homogenous polygons, which often results in poor results due to lack of spectral variability. Within each satellite data scene, the training data was generated approximately proportional to the expected land cover types for that scene. This method of training selection has been found to increase accuracy of resulting classifications (Zhu et al. 2016).

The training data were refined through an iterative process of collection, test classification, and accuracy assessment. This process is used to ensure each land cover type in a satellite image is accurately represented and the classification models do not overfit. In total, more than 30,000 pixels of training data were extracted for each classifier. However, the final training datasets were randomly and proportionally reduced to about 8,000 pixels in order to maintain operational capability of the Python-based tool for end-users.

A Python 3 script was written to read in the training data for further processing and to train the classifiers. The data were processed primarily using NumPy (version 1.16.5) and Scikit-learn (version 0.22.1), assisted by Geospatial Data Abstraction Library (GDAL) (version 2.3.3) for image processing and joblib (version 0.15.1) for parallelization of code for increased speed.

For each of the semiautomatic land cover classification models, the training data are split into training and testing samples: 80% of the data are used for training, and the remaining 20% are withheld and used for model evaluation and initial accuracy assessment. The proportions of the training data are maintained in the split.

For each of the three models, two-date Sentinel-2 classifier, single-date Sentinel-2 classifier, and two-date Landsat 8 classifier separate classifiers were created with the corresponding training data. An ensemble of three SVMs were built using Scikit-learn version 0.22.1 using the C-based SVM

built on Libsvm with a linear kernel, regularization parameter of 1.0, balanced class weights, cache size of 750mb, and all other parameters as default. The ensemble was created through the "sklearn.ensemble.baggingclassifier" function, which is a meta-estimator that fits random subsets of the training data to create three separate SVM classifiers. The algorithm then aggregates the individual classifier predictions to determine the final prediction based on majority vote. This method reduces classifier variance and can lead to a more robust model than a single classifier (Breiman 2001).

The trained models are then exported from the Python environment by Python pickle serialization and thus persist for later use in the custom ArcGIS Pro Python toolbox.

This custom ArcGIS Pro toolbox would be a part of a larger suite of tools that support terrain analysis applications. While current capabilities rely on outdated land cover products or require manual creation of training data, this method with an easy-to-use interface will enable on-demand mapping from the desired time period with minimal user input. As a result, the tool will provide up-to-date land cover maps at multiple spatial resolutions in order to enable timely, up-to-date, and accurate decision-making cababilities.

3.2 Python toolbox design

3.2.1 ArcGIS Pro Python toolbox background

ArcGIS Pro contains a suite of commercial, closed-source geospatial processing tools. These tools can be accessed within the ArcGIS Pro Graphical User Interface (GUI) through the ArcToolBox functionality. If custom functionality integrating third-party Python libraries is desired, then custom Python toolboxes must be built.

A Python toolbox can be directly integrated into the collection of existing geoprocessing tools found in ArcToolBox. The difference is that the Python toolbox must be created entirely in Python based on object-oriented programming structure and a toolbox template. With some effort, Python toolboxes can leverage third-party scientific Python libraries directly within ArcGIS Pro and thereby provide advanced functionality for custom applications visualized through the ArcGIS Pro GUI.

Python toolboxes in ArcGIS have some limitations that make them difficult to work with. The Python toolboxes are finicky and prone to various errors. Multithreading capabilities are severely limited. For example, in order to use parallel processing the toolbox must be meticulously structured with (1) a single Python toolbox and (2) separate Python scripts that serve as importable modules. Further, the inherent parallel processing of joblib and Scikit-learn libraries is incompatible. The Python toolbox designer must rewrite their code structure and function using the more cumbersome multiprocessing Python library.

3.2.2 ETP Land Cover mapping Python toolboxes

A set of image preprocessing tools and cloud masking tools are included in the custom Python toolbox as shown in Figure 4. The main functionality of the custom ETP Land Cover toolbox is the three land cover type classification tools created based on single-date Sentinel-2, two-date Sentinel-2, and two-date Landsat 8 classification models. The single-date Sentinel-2 classification tool produces a 6-class land cover type map, whereas both of the two-date classification tools produce 11-class land cover type maps. The classes represented in each classification tool are shown in Table 2.

Figure 4. View of the land cover tools as shown integrated into ArcGIS Pro as a Python toolbox. The tools are integrated as a subset of the Enhanced Terrain Processing Toolkit.

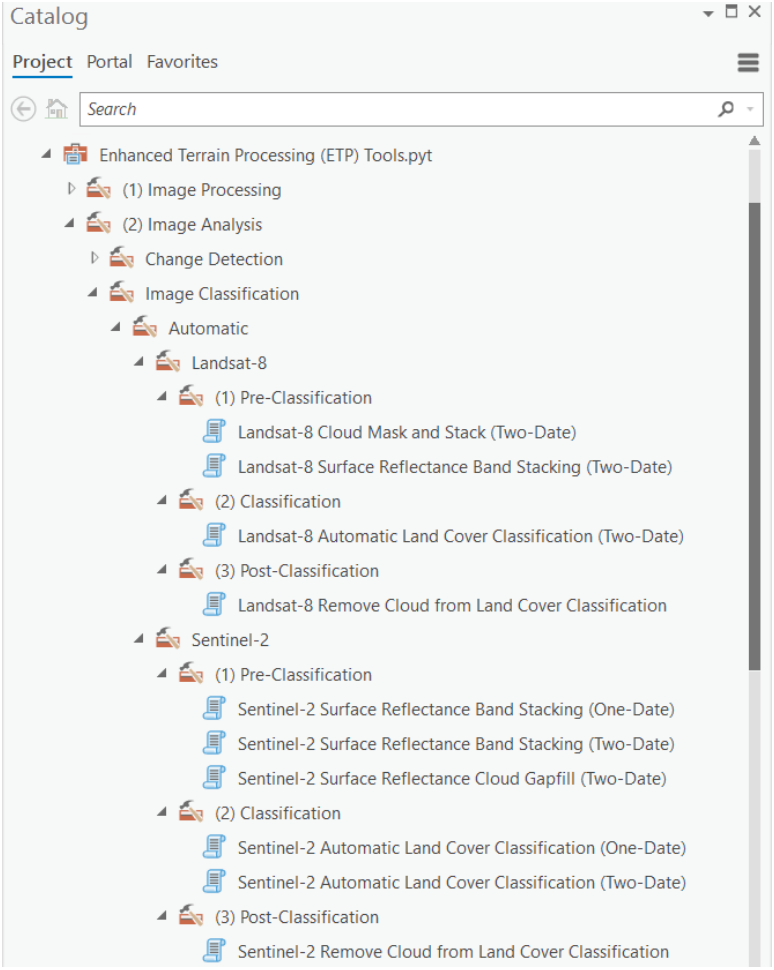


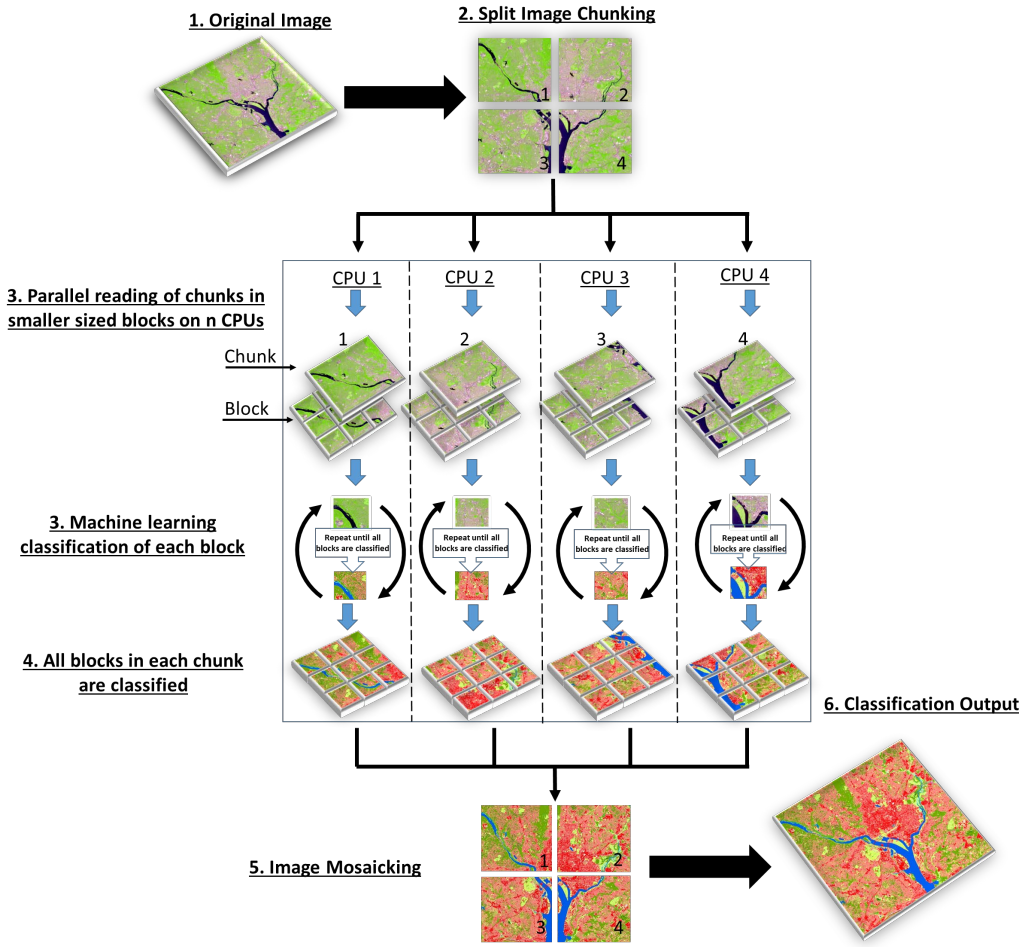
Table 2. Land cover types generated by the single date and two date classification tools.

Single-Date Classification Tool Land Cover Types	Two-Date Classification Tool Land Cover Types
Water	Water
Barren	Barren
Scrub, Shrub, Grasses	Scrub, Shrub, Grasses
Built-up Area	Deciduous Trees
Wetlands	Evergreen Trees
Trees	Herbaceous Wetlands
-	Woody Wetlands
-	Perennial Water
-	Built-up (Low Density)
-	Built-up (High Density)
-	Cropland

In order to enable the land cover classification workflow in ArcGIS Pro, we develop an ArcGIS Pro Python toolbox. The Python toolbox works in the same manner as a standard ArcGIS tool but is entirely written in Python and can accommodate custom capabilities unavailable in ArcGIS and third party Python libraries.

The Python toolbox is designed to include image chunking and parallel processing in order to enable faster data processing and processing on computers with reduced computing capability. The general process works the same for each of the three tools, just with different input imagery and models. First, Sentinel-2 or Landsat 8 imagery is input into the tool, and the imagery is split into equally sized fractions (chunks) equal to the number of central processing units (CPUs) on the user's machine. Each fraction of an image, or image chunk, is then run in parallel on the user's machine to predict the land cover type as shown in Figure 5. The image chunks on each CPU are read only partially into memory at once using a block-by-block reading and writing method. In this case, each image chunk is read and written in 400 × 400 pixel blocks. This procedure saves computer memory usage by reading smaller portions of the image instead of all at once, which can be memory intensive. Once parallel processing of all image chunks is completed, the image chunks are mosaicked together at the end to combine the output back into a single file. Overall, this process saves time by running multiple processing tasks at once through parallelization.

Figure 5. Parallel processing and image chunking process that enables faster computation and compatibility on systems with limited memory within ArcGIS Pro tools.



3.3 Training data

Training data in machine learning is the backbone of a successful algorithm and accurate classification. This data is used by the algorithm to recognize and understand patterns and create information that it will use for predictions. Figure 6 represents an example of three land cover types used to create training data. It can be seen that the pixel distribution within each of the training samples is uniform as to equally represent the signal within each class. This training data is extracted from the multispectral Sentinel-2 and Landsat 8 image stacks shown in Figure 6. Since both satellite sensors have different spatial and spectral resolutions, a unique training dataset is built for each land cover classification tool. This training data is a critical part of the classification tool, thus each class is equally represented by the training samples and all pixels were uniformly selected and representative of the entire landscapes.

Imagery from both sensors was used for each site so as to avoid bias when generating training data and training the classification models. For each dataset and site, corresponding winter (leaf-off) and summer (leaf-on) images were downloaded and used to build the training data for the two date classifiers, whereas summer (leaf-on) imagery was used for the Sentinel-2, single-date classifier.

Figure 6. Example of three land cover types used as training data for the land cover classification tools.

3.4 Model training

Once training data is created, an ensemble of three linear SVM models is built and 80% of the training data is applied. The remaining 20% is withheld for accuracy assessment. The SVM is fit in order to enable the model to classify land cover type. The model is trained using Python and the SVM function with Scikit-learn. Separate models were created for the Landsat 8 two-date classifier, Sentinel-2 two-date classifier, and Sentinel-2 single-date classifier. The trained models are then transferred and written into ArcGIS Pro Python toolboxes to enable the prediction of land cover type across different regions based on user input. These trained models can later be updated based on improvements to the model, such as expansion to outside CONUS areas.

4 Example Output and Model Evaluation

4.1 Example output

The user can create an output land cover type map by acquiring Sentinel-2 or Landsat 8 imagery in the surface reflectance format (ARD for Landsat 8, L2A for Sentinel-2) and by using the ETP toolbox preprocessing tools to composite the imagery. The land cover map can be produced by running the respective tool of Landsat 8 two date, Sentinel-2 two date, or Sentinel-2 single date. The total time to produce a land cover map depends on the user's machine, but for a laptop with an 8 core CPU and 16GB memory, the process took about 20 minutes total, including preprocessing.

Figure 7 shows example output for the Sentinel-2 two-date classification tool over three different locations. Visually comparing the tool output with respect to the original Sentinel-2 images, we can see that across all three test sites, the water class, wetlands, and built-up areas (both high and low intensity) are well classified by the tool. For example, open water bodies as seen in the Savannah, Georgia site as well as smaller rivers and lakes as shown in Minneapolis and Washington, DC are well defined by the classification tool. Specifically, the distinction between wetland and water features can be seen in the Savannah, Georgia site. Wetland classification is a challenging problem as the spectral information can easily be confused with other land cover classes due to the lack of a single, unifying land cover feature as well as wetlands' highly dynamic nature (e.g., variation in water levels) causing confusion between the mixed pixels and surrounding areas. Nevertheless, by using two images from different dates, the algorithm can utilize the changes in spectral differences between the scenes and aid in the identification of wetlands in comparison to less dynamic land cover types that may be more uniform (e.g., water, bare areas). Similarly, this is also captured when mapping high-intensity versus low-intensity built-up areas. As seen by the Washington, DC, example, the classification tool is able to distinguish between the densely downtown areas of Washington, DC, in comparison to the suburban regions surrounding the city.

Figure 7. Example output for the Sentinel-2 two-date classifier over three different sites.

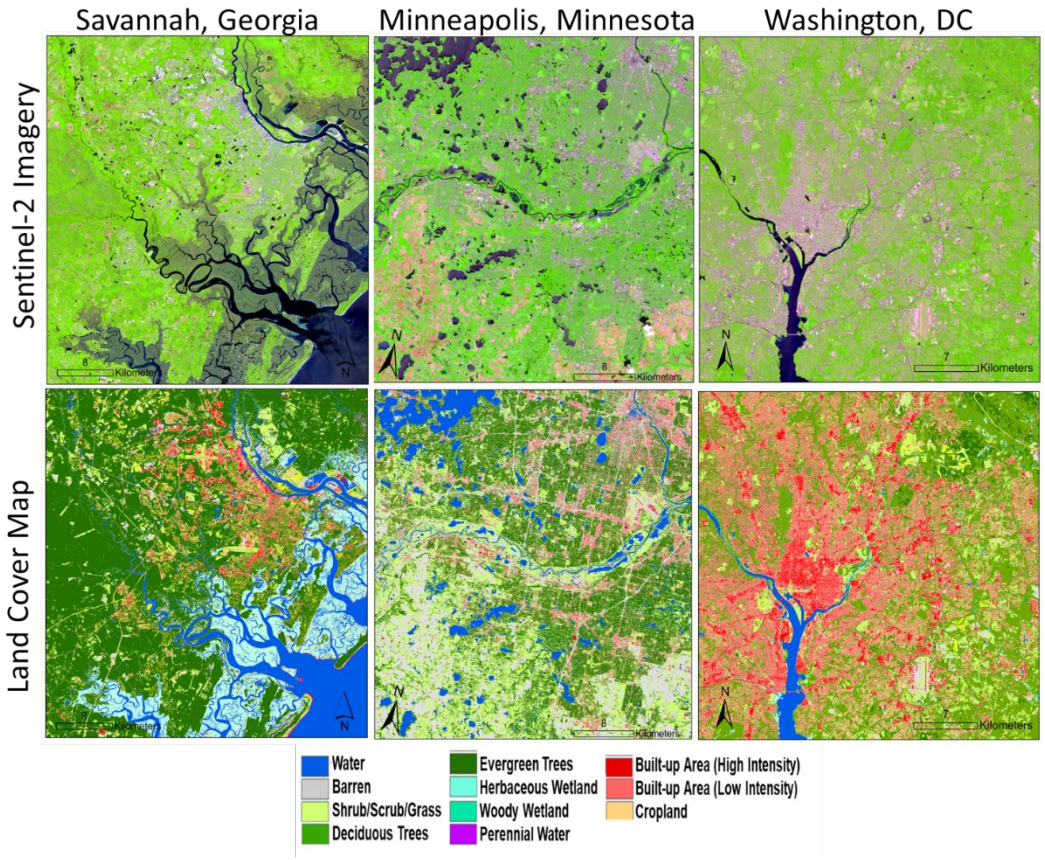
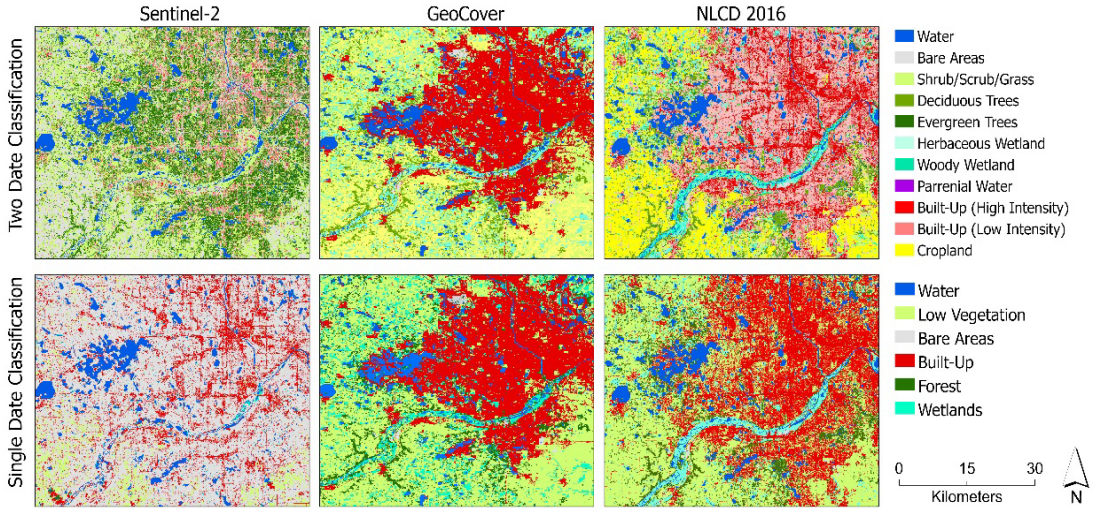


Figure 8 focuses on the Minneapolis test site and shows a comparison between the Sentinel-2 single-date and two-date classification land cover output with respect to one another as well as to the 30m GeoCover and 2016 NLCD. Comparing the two-date classification output with the single-date classification, a clear distinction occurs in the bare areas (single date) versus shrub/scrub/grass land cover type (two-date). The single-date classifier identifies the land cover makeup for the area to be mostly bare, whereas the two-date classifier identifies the region as mostly low vegetation with deciduous and evergreen trees. Additionally, both tools show similar spatial patterns of urban areas; however, many of the mixed pixels in the urban areas are classified as bare area in the single-date classifier.

Figure 8. Comparison of Sentinel-2 two-date and single-date classification with reclassified GeoCover and 2016 NLCD maps over Minneapolis, MN.

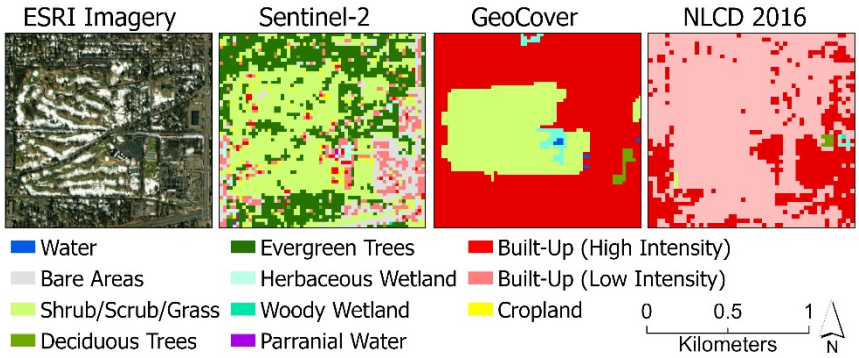


Comparing the two-date and single-date classification output to the 30m GeoCover and NLCD 2016, we can see that the water and wetland classes are most consistent across all comparisons, followed by the urban land cover class. Both GeoCover and NLCD 2016 show a highly dense land cover type for Minneapolis, whereas the Sentinel-2 classification product shows a built-up land cover type with a mix of shrub/scrub/grass and evergreen trees. This is due to the way the way the different models are trained, and it appears that the NLCD and GeoCover product tend to label mixed pixels as built-up area. This is not to say that either the Sentinel-2 product or the NLCD/GeoCover is technically wrong, only that inherently they are favoring a more dominant class. For example, the Sentinel-2 tool incorporates the dynamic changes in vegetation/tree cover over the two scenes. It is important to note that the spectral and spatial resolution of Sentinel-2, GeoCover, and NLCD 2016 are all different, which can also impact the result. NLCD and GeoCover derive the land cover product from a dense time series of image composites, whereas both the Sentinel-2 land cover tools are based on either a single image (summer/single-date classifier) or on leaf-on and leaf-off images (Sentinel-2 two-date classifier).

Figure 9 represents an example of an instance where the Sentinel-2 land cover product is showing an agreement of land cover type with the GeoCover product but a different land cover type than the NLDC product. The example shows a golf course marked as a shrub/scrub/grass land cover type both within the two-date Sentinel-2 and GeoCover, whereas the NLCD identifies that same golf course as a low-intensity built-up area. It

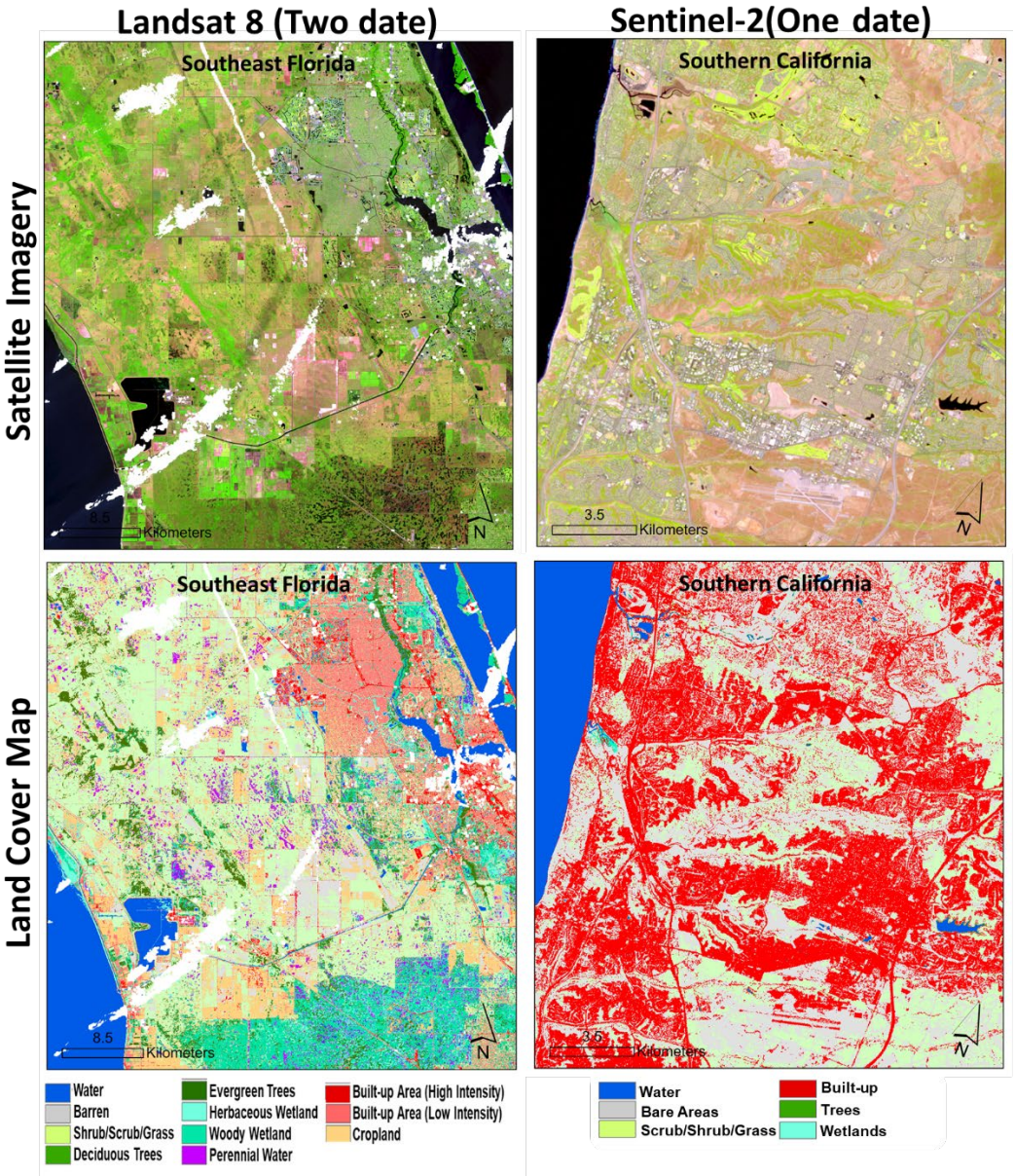
appears the NLCD product is classifying land use whereas the other two products are classifying land cover. This land use versus land cover distinction also likely holds true for other land cover classes in addition to built-up. For comparison reasons, we have also provided an ESRI imagery basemap layer shown in red, green, blue (RGB) as a reference.

Figure 9. Comparison of Sentinel-2 two-date classifier with GeoCover and NLCD products over a golf course, highlighting the differences between land use and land cover reflected among the classifiers.



The Sentinel-2 two-date tool is preferred over the others due to higher spatial resolution than Landsat 8, additional land cover classes than the single date, as well as more spatially consistent performance than the single date. Example output for the Landsat 8 two-date tool and Sentinel-2 single-date tool are shown in Figure 10.

Figure 10. Landsat 8 (two date) and Sentinel-2 (single date) imagery and resulting land cover map over South Florida and Southern California respectively.



4.2 Accuracy assessment

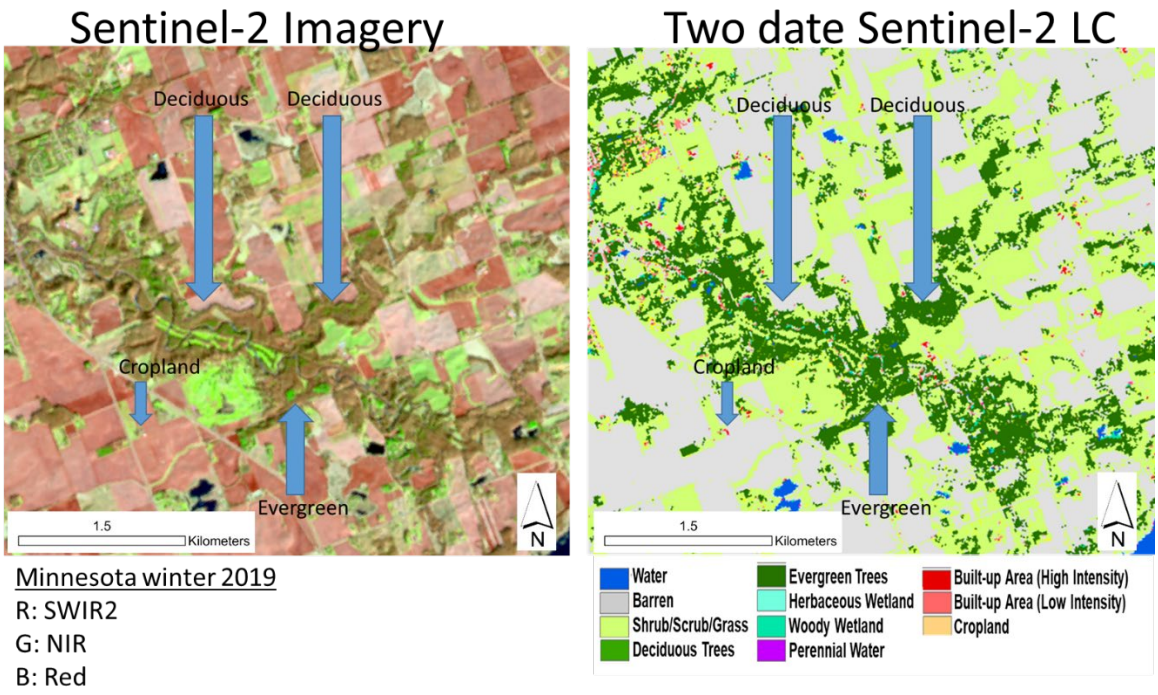
The accuracy assessment is a critical part of any image classification project, and a project should not be considered complete without it. To determine an initial assessment of accuracy, we withheld 20% of the samples from the training data at random and proportional to each class. This withheld data was not used to train each classifier but was used only for model evaluation. This evaluation provides insight into the accuracy of the models. Note that accuracy across different landscapes and terrain are

likely to vary based on the values shown in Table 3 and Table 4 because these do not include additional points from a test dataset, which will be conducted in an updated version of our algorithm in a subsequent report. These variations could be better quantified through a more time-consuming accuracy assessment that is beyond the scope of this report, as the models are undergoing improvements annually. The accuracy metrics for the Sentinel-2 single-date classifier, Sentinel-2 two-date classifier, and Landsat 8 two-date classifier are shown in Table 3 and Table 4.

The overall average accuracy for Sentinel-2 single-date classifier is about 85% as compared with about 82% for the Sentinel-2 two-date classifier. These metrics exceed the overall goal of 80% or greater accuracy as dictated by the mission requirements. There are several reasons why the Sentinel-2 single-date classifier accuracy was slightly higher. One reason is attributed to lower class precision; it is much easier to map 6 classes than it is to map 11 classes. The single-date classifier only needs to determine if a pixel is tree, whereas for the two date classier it needs more precision and must specify if this pixel is evergreen tree or deciduous tree. Thus, a lower accuracy is expected. It is also likely that some errors can be attributed to misclassifications like this example.

Of the different classes for the two-date Sentinel-2 classifier, we found relatively lower accuracies for both built-up categories, woody wetlands, bare areas, and low vegetation. Through development and testing we found that woody wetlands were sometimes confused with other tree categories, while bare areas and built-up areas were often confused with each other due to relatively similar spectral signatures, especially across semi-arid landscapes. Through testing, we also observed confusion between evergreen and deciduous trees in some testing locations. Figure 11 shows a large patch of mostly deciduous trees that were classified as evergreen trees. After looking at the input training data, we noticed that evergreen trees in some regions (e.g., from southern US and western US) can resemble the spectral signal of deciduous trees in other locations which can make class separability difficult. Figure 11 also shows an area that is clearly a cropland land use type, but in both dates of imagery the land cover type was bare. Thus, this is not an error in the classifier as it is merely a difference of land use and land cover. It is worth pointing out as this classifier captures land cover, not land use.

Figure 11. Example of Sentinel-2 image and associated land cover type output showing misclassification of evergreen and deciduous trees.



We found the highest accuracies for the water, deciduous trees, evergreen trees, and herbaceous wetland categories with F1-scores of 99%, 90%, 89%, and 93% respectively. However, we caution that F1-scores will vary across different locations and likely be lower in reality, especially once compared against a separate validation dataset.

Some of the error is attributed to heterogeneous pixels containing more than one land cover type or not containing a clear majority land cover type, thereby making accurate prediction difficult. For example, tree cover that partially obstructs a built-up surface such as a road. This could be classified as either tree cover or built-up. Neither classification would be wrong, but neither would be fully accurate either. These types of issues also likely explain some of the reported lower precision and recall for built-up classes and scrub/shrub/grass.

The Landsat 8 two-date classifier results are shown in Table 3. The accuracy, as expected, is relatively similar to the other classifiers with 86% overall accuracy. Compared with the Sentinel-2 two-date classifier, the Landsat 8 two-date classifier appeared to perform better in several categories such as low vegetation and woody wetlands. The Landsat 8 classifier performed worse on dynamic vegetation (cropland)–likely due to

fewer near-infrared bands, which can be helpful discrimination of different vegetation types. Further, we note the accuracies are somewhat different because different training data were collected for each of the classifiers, and we note that slightly more training data used was in the Landsat 8 classifier.

Table 3. Accuracy metrics for the Sentinel-2 and Landsat 8 two-date classifier for each of the 11 land cover classes. Precision, recall, and F1-scores are shown with pixel count analyzed for each class.

	Sentinel-2				Landsat-8			
Class	Precision	Recall	F1-Score	Pixel Count	Precision	Recall	F1-Score	Pixel Count
Water	0.98	0.99	0.99	203	0.98	0.99	0.99	448
Bare Areas	0.65	0.90	0.75	142	0.71	0.92	0.82	169
Low Vegetation	0.78	0.76	0.77	339	0.84	0.91	0.88	369
Deciduous Trees	0.89	0.92	0.90	237	0.90	0.94	0.92	346
Evergreen Trees	0.91	0.86	0.89	236	0.92	0.90	0.91	450
Herbaceous Wetlands	0.97	0.90	0.93	125	0.94	0.89	0.92	149
Woody Wetlands	0.66	0.77	0.72	35	0.81	0.80	0.81	192
Temporary Water	0.94	0.79	0.86	28	0.80	0.91	0.86	89
Built-Up	0.71	0.77	0.74	143	0.75	0.87	0.81	217
Build-Up (Low Intensity)	0.73	0.69	0.71	192	0.81	0.77	0.79	203
Dynamic Vegetation	0.83	0.80	0.81	288	0.77	0.74	0.76	312
Overall Accuracy	0.82				0.86			

Table 4. Accuracy metrics for the Sentinel-2 single-date classifier for each of the 6 land cover classes. Precision, recall, and F1-scores are shown with pixel count analyzed for each class.

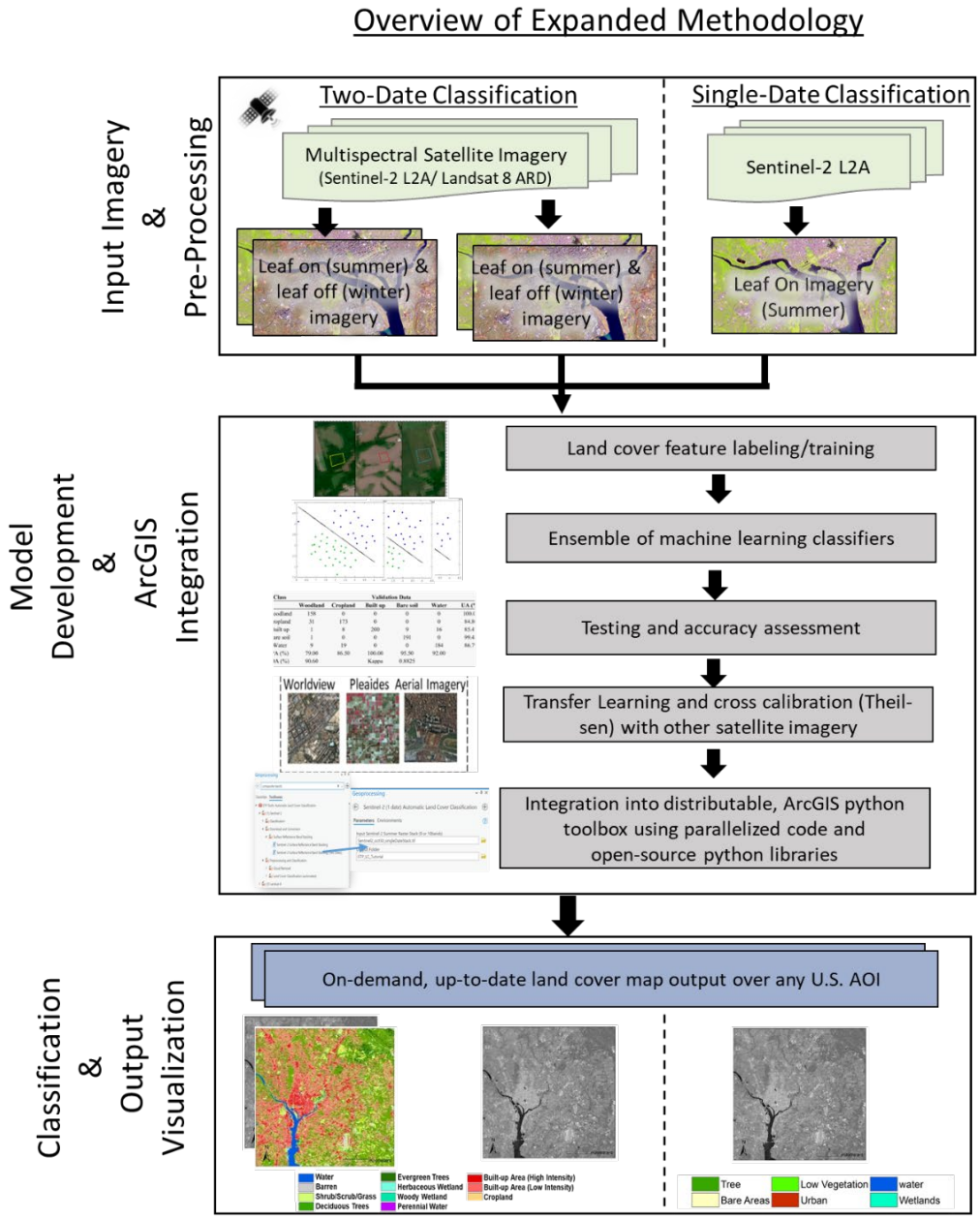
Class	Precision	Recall	F1-Score	Pixel Count
Water	0.97	0.97	0.97	715
Bare Areas	0.86	0.79	0.82	219
Low Vegetation	0.80	0.91	0.85	497
Trees	0.92	0.88	0.90	618
Wetlands	0.81	0.84	0.83	110
Built-Up	0.72	0.77	0.75	277
Overall Accuracy	0.85			

While the land cover mapping tools can likely be used on imagery outside of CONUS, the tools have not been evaluated or designed for this purpose. It is likely that reduced accuracy would be found if a user applies the tool to Outside Continental United States (OCONUS) imagery because training data was not yet collected there (future additions will include OCONUS imagery compatibility). Further, the single-date classifier relies on imagery from leaf-on season (e.g., not winter), while the two-date classifiers rely on snow-free imagery, one from leaf-on (summer/spring) and one from leaf-off (winter) time periods. Any pixels containing snow will result in misclassification as a different category such as built-up because there is not yet a snow category. Future iterations of the model plan to address the snow issue by including a snow cover class, as well as expanding the tools to include OCONUS areas.

4.3 Advancements with synthetic training data for sensor agnostic mapping

As mentioned in the methodology section, the ultimate goal is to create an automated land cover type classification model and associated decision support tools to execute the model for an end user on more imagery than just Landsat 8 and Sentinel-2. Such tools allow the end user to generate up-to-date land cover type maps at higher temporal and spatial resolutions in comparison to outdated and coarser land cover products. Figure 12 highlights the slightly modified methodology used to achieve this goal.

Figure 12. Expanded methodology for sensor-agnostic approach using Theil-Sen regression to transform training data of Sentinel-2 to represent synthetic data for the target sensor.



Sentinel-2 and Landsat 8 imagery are used to create training data for selected land cover types across different ecological regions within CONUS. This data is then fed into a robust regression using Theil-Sen to transform the existing Sentinel-2 training data into synthetic training data representative of the target sensor, such as WorldView-3, Pleaides, or other airborne assets. The linear transformation from the Theil-Sen regression will be based on each Sentinel-2 band and corresponding, similar spectral bands from other sensors. Imagery will be aggregated and aligned to the same pixel size to match Sentinel-2. The resulting linear

relationships can be applied to the Sentinel-2 training data and can convert it into synthetic training data for the other sensor.

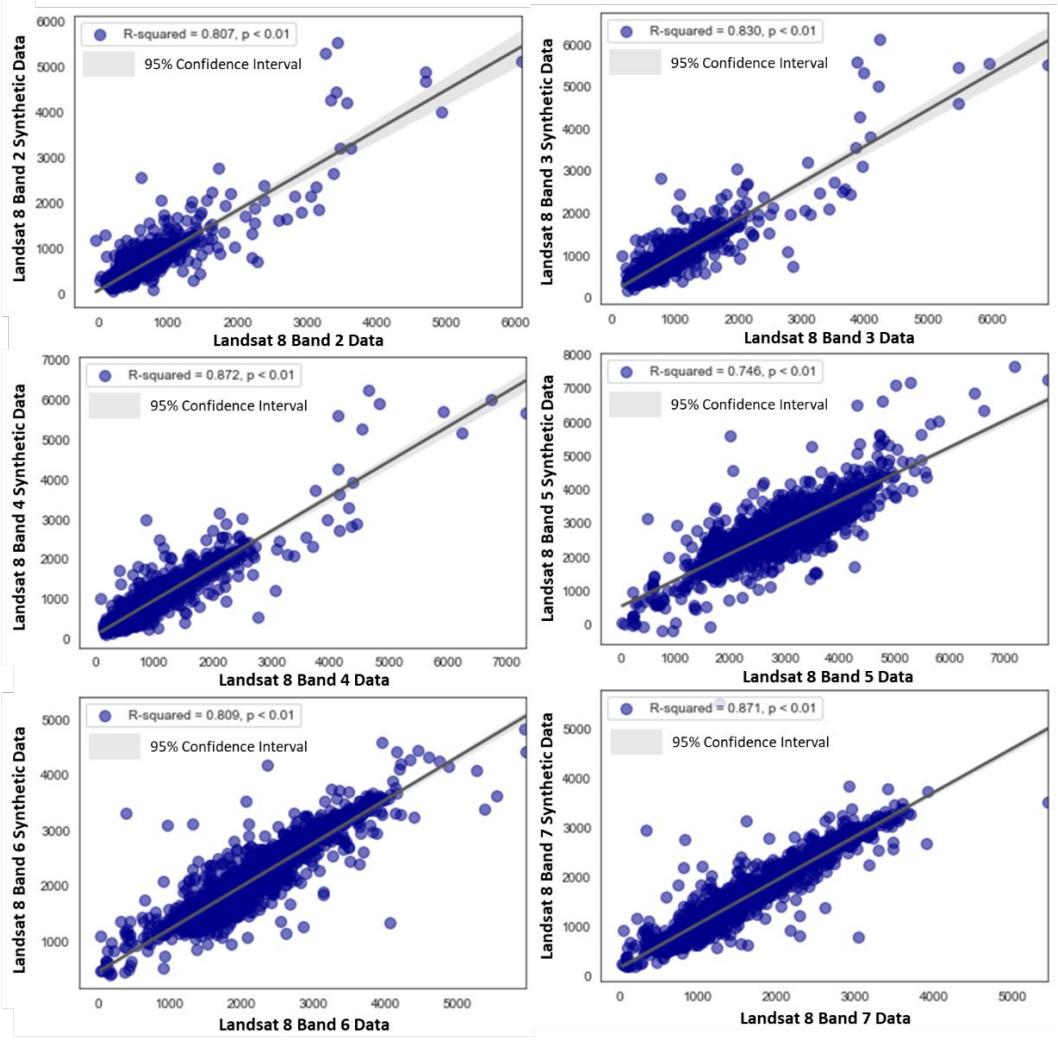
Ultimately, this sensor-agnostic model would be wrapped into a Python-based ArcGIS Pro toolbox to support rapid land cover type mapping for multispectral satellite and aerial imagery by end users. This would enable use for more sensors than just Landsat 8 or Sentinel-2; it could also include digital globe assets and more without the need to manually generate additional training data.

There is potential to generate synthetic training data for Landsat 8 using the training data created with Sentinel-2 polygons and Sentinel-2 spectral values. This same methodology could be applied to other imagery assets of higher spatial resolution, such as Worldview-3. The generation of synthetic training data would save significant time without the need to manually label pixels for every given sensor.

For prototyping purposes, we acquired coincident cloud-masked Landsat 8 and Sentinel-2 imagery from the same acquisition date for Washington, DC, Los Alamos, New Mexico, and northern California, representing three unique climates. Both sets of imagery were resampled to 60m spatial resolution in order to align pixels to the lowest common spatial resolution of the two sensors. To get a spectrally balanced sample from each scene, we collected 600 points in each scene generated by stratified random sampling. The stratification was based on equal area sampling of the Normalized-Difference Vegetation Index (NDVI) with 6 equal interval classes (e.g., NDVI 1–0.66, 0.66–0.33, 0.33–0, etc.). Theil-Sen regression was performed between Sentinel-2 bands and Landsat 8 bands in order to generate a linear relationship to transform our Sentinel-2 training data into synthetic Landsat 8 training data. Unlike ordinary least squares regression, Theil-Sen is robust and accounts for the offsetting effects of outliers. The spectral bands are slightly different between Landsat 8 and Sentinel-2, so we regressed the most similar bands together (i.e., Landsat 8 Band 2 with Sentinel-2 band 1, Landsat 8 band 3 with Sentinel-2 band 2, Landsat 8 band 4 with Sentinel-2 band 3, Landsat 8 band 5 with Sentinel-2 band 7, and Landsat 8 band 6 with Sentinel-2 band 8). The process of generating synthetic training data for Landsat 8 is used to show how synthetically generated data could save time for applying the classifier to other sensor types.

Figure 13 shows the relationship between each Landsat 8 band and the synthetically generated Landsat 8 data based on Theil-Sen regression between the Sentinel-2 and Landsat 8 imagery. Overall, the relationships are strong; each band has R-squared exceeding 0.74. This suggests that transforming the Sentinel-2 training data based on this relationship may lead to accurate models for Landsat 8. This relationship could easily be applied to generate training data and classification models for other high resolution sensors such as Worldview.

Figure 13. Relationship of Landsat 8 data to synthetically generated Landsat 8 data shown for DC, New Mexico, and northern California data points.



Theil-Sen regression indicated promising results. Table 5 shows the root mean square error (RMSE) for the Theil-Sen regression of each Sentinel-2, Landsat 8 band combination across the three sites (New Mexico, California, DC) and compared with a single site (California). In each case and as expected, the synthetically generated Landsat 8 imagery (labeled as

"predicted") had a lower RMSE than the original Landsat 8 imagery did as compared with the Sentinel-2 imagery for each band. We show the RMSE for the single California site to demonstrate that model error is further reduced if the Theil-Sen regression is applied using location-specific data. RMSE averaged 165 reflectance units (single site) versus 294 reflectance units (three site). From the Theil-Sen regression, coefficients and intercept can be extracted and applied direcly to the training data for transformation into synthetic data. This result suggests that synthetic training data accuracy would be best by using local site imagery only for transforming the Sentinel-2 data into synthetic training data for the sensor of interest. Further testing and exploration should be conducted to evaluate the results from training classifiers with this synthetic data.

Table 5. Relationship of Landsat 8 data to synthetically generated Landsat 8 data shown for DC, New Mexico, and northern California data points.

Band Combination	RMSE S2, actual L8 (1 image)	RMSE S2, predicted L8 (1 image)	RMSE S2, actual L8 (3 images)	RMSE S2, predicted L8 (3 images)
S2 Band 2, L8 Band 2	241	144	354	204
S2 Band 3, L8 Band 3	233	155	342	223
S2 Band 4, L8 Band 4	259	161	377	251
S2 Band 8A, L8 Band 5	395	229	522	491
S2 Band 11, L8 Band 6	388	174	486	328
S2 Band 12, L8 Band 7	322	131	453	266

4.4 Installation and usage of ArcGIS Pro Tools

The land cover mapping capabilities have been written into a custom ArcGIS Pro Python-based toolbox compatible with any windows computer with ArcGIS Pro version 2.5 or higher and 8GB RAM or more. Easy-to-follow instructions on installation, usage, and use cases are available in separate tutorial and user guide documents.

Conclusion

This project led the development of pretrained machine learning classifiers to create land cover type models and associated decision support tools for use in the CONUS region. The pretrained machine learning models were created for use with a single date of Sentinel-2 surface reflectance imagery as well as two dates of Sentinel-2 or two dates of Landsat 8. Each of the three separate satellite imagery models were then integrated into ArcGIS Pro GUI using Python toolboxes with parallel processing and image chunking to serve as easily accessible decision support tools for U.S. Army Corps of Engineers (USACE) customers. The decision support tools enable a user to create a land cover type map without the need to generate training data or to rely on costly or outdated static maps. Further, the ArcGIS Pro tools enable a user to generate a map in a matter of minutes with minimal user input. The land cover type classes can easily be recategorized for interchangeability with other land cover maps used in terrain analysis applications such as VizNav.

The estimated overall accuracy of the three pre-trained models is 82% (Sentinel-2 two date), 86% (Landsat 8 two date), and 85% (Sentinel-2 one date), which suggests generally good performance across most land cover types. After testing the imagery across varied ecoregions of CONUS, the study found future areas of improvement to the models should include better separability between evergreen and deciduous trees, as well as the separability between bare areas and built-up areas. These errors are attributed to spectral similarities found across the geographically distributed training imagery.

Ultimately, the semi-automated land cover mapping tools serve as a foundation on which to build more advanced land cover mapping tools. We also showcased a methodology that will leverage training data built on one sensor platform and apply Theil-Sen robust regression to generate synthetic training data for other imagery platforms. This would expand the sensor capabilities and add additional models.

Future improvements under this project will expand the geographic coverage of the models outside of CONUS, incorporate advanced deep learning models for more robust classification, account for snow cover in order to increase versatility, and expand into higher spatial resolution sensors.

References

Breiman, L. 2001. “Random Forests.” *Machine Learning 45* (1): 5–32.

Büttner, G. 2014. “CORINE Land Cover and Land Cover Change Products. *Land Use and Land Cover Mapping in Europe*. Dordrecht: Springer, 55–74.

Congalton, R. G., and K. Green. 2019. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, FL: CRC press.

Cunningham, D., J. E. Melican, E. Wemmelmann, and T. B. Jones. 2002. “GeoCover LC—A Moderate Resolution Global Land Cover Database. In *ESRI International User Conference* (July).

Foody, G. M., and A. Mathur. “Toward Intelligent Training of Supervised Image Classifications: Directing Training Data Acquisition for SVM Classification. *Remote Sensing of Environment* 93 (1–2): 107–17.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, and A. Kommareddy. 2013. “High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342 (6160): 850–53.

Hasmadi, M., H. A. Pakhriazad, and M. F. Shahrin. 2009. “Evaluating Supervised and Unsupervised Techniques for Land Cover Mapping Using Remote Sensing Data.” *Geografia: Malaysian Journal of Society and Space* 5 (1): 1–10.

Jia, K., Q. Li, Y. Tian, B. Wu, F. Zhang, and J. Meng. 2012. “Crop Classification Using Multi-Configuration SAR Data in the North China Plain.” *International Journal of Remote Sensing* 33 (1): 170–83.

Jin, S., C. Homer, L. Yang, P. Danielson, J. Dewitz, C. Li, Z. Zhu, G. Xian, and D. Howard. 2019. “Overall Methodology Design for the United States National Land Cover Database 2016 Products.” *Remote Sensing* 11 (24): 2971.

Kontgis, C., A. Schneider, and M. Ozdogan. 2015. “Mapping Rice Paddy Extent and Intensification in the Vietnamese Mekong River Delta with Dense Time Stacks of Landsat Data.” *Remote Sensing of Environment* 169: 255–69.

Kussul, N., M. Lavreniuk, S. Skakun, and A. Shelestov. 2017. “Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data.” *IEEE Geoscience and Remote Sensing Letters* 14 (5): 778–82.

Lasko, K., K. P. Vadrevu, V. T. Tran, and C. Justice. 2018. “Mapping Double and Single Crop Paddy Rice with Sentinel-1A at Varying Spatial Scales and Polarizations in Hanoi, Vietnam.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11 (2): 498–512.

Lavreniuk, M., N. Kussul, S. Skakun, A. Shelestov, and B. Yailymov. 2015. “Regional Retrospective High Resolution Land Cover for Ukraine: Methodology and Results.” *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (July): 3965–68. IEEE.

Malinowski, R., S. Lewiński, M. Rybicki, E. Gromny, M. Jenerowicz, M. Krupiński, A. Nowakowski, C. Wojtkowski, M. Krupiński, E. Krätzschmar, and P. Schauer. 2020. “Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery.” *Remote Sensing* 12 (21): 3523.

Olofsson, P., G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder. 2014. “Good Practices for Estimating Area and Assessing Accuracy of Land Change.” *Remote Sensing of Environment* 148: 42–57.

Pal, M., and G. M. Foody. 2012. “Evaluation of SVM, RVM and SMLR for Accurate Image Classification with Limited Ground Data.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5 (5): 1344–55.

Pal, M., and P. M. Mather. 2005. “Support Vector Machines for Classification in Remote Sensing.” *International Journal of Remote Sensing* 26 (5): 1007–11.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and J. Vanderplas. 2011. “Scikit-Learn: Machine Learning in Python.” *The Journal of Machine Learning Research* 12: 2825–30.

Pickens, A. H., M. C. Hansen, M. Hancher, S. V. Stehman, A. Tyukavina, P. Potapov, B. Marroquin, and Z. Sherani. 2020. “Mapping and Sampling to Characterize Global Inland Water Dynamics from 1999 to 2018 with Full Landsat Time-Series.” *Remote Sensing of Environment* 243: 111792.

Suykens, J. A., and J. Vandewalle. 1999. “Least Squares Support Vector Machine Classifiers.” *Neural Processing Letters* 9 (3): 293–300.

Teluguntla, P., P. S. Thenkabail, A. Oliphant, J. Xiong, M. K. Gumma, R. G. Congalton, K. Yadav, and A. Huete. 2018. “A 30-m Landsat-Derived Cropland Extent Product of Australia and China Using Random Forest Machine Learning Algorithm on Google Earth Engine Cloud Computing Platform.” *ISPRS Journal of Photogrammetry and Remote Sensing* 144: 325–40.

Xiong, J., P. S. Thenkabail, M. K. Gumma, P. Teluguntla, J. Poehnelt, R. G. Congalton, K. Yadav, and D. Thau. 2017. “Automated Cropland Mapping of Continental Africa Using Google Earth Engine Cloud Computing. *ISPRS Journal of Photogrammetry and Remote Sensing* 126: 225–44.

Zhu, Z., A. L. Gallant, C. E. Woodcock, B. Pengra,P. Olofsson, T. R. Loveland, S. Jin, D. Dahal, L. Yang, and R. F. Auch. 2016. “Optimizing Selection of Training and Auxiliary Data for Operational Land Cover Classification for the LCMAP Initiative. *ISPRS Journal of Photogrammetry and Remote Sensing* 122: 206–22.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
November 2021	Final report	FY20–FY21

4. TITLE AND SUBTITLE

Semi-Automated Land Cover Mapping Using an Ensemble of Support Vector Machines with Moderate Resolution Imagery Integrated into a Custom Decision Support Tool

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
633463

6. AUTHOR(S)

Dr. Kristofer Lasko, Dr. Elena Sava

5d. PROJECT NUMBER
AU1

5e. TASK NUMBER
SAU102

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Geospatial Research Laboratory
U.S. Army Engineer Research and Development Center
7701 Telegraph Road
Alexandria, VA 22315-3864

8. PERFORMING ORGANIZATION REPORT NUMBER

ERDC/GRL-TR-21-7

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Geospatial Research Laboratory
U.S. Army Engineer Research and Development Center
7701 Telegraph Road
Alexandria, VA 22315-3864

10. SPONSOR/MONITOR'S ACRONYM(S)

ERDC-GRL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES Tactical Geospatial Information Capabilities, Enhanced Terrain Processing—Demonstration

14. ABSTRACT

Land cover type is a fundamental remote sensing–derived variable for terrain analysis and environmental mapping applications. The currently available products are produced only for a single season or an entire year. These products have a coarse resolution and quickly become outdated, as land cover type can undergo significant change over a short time period. In order to enable on-demand generation of timely and accurate land cover type products, we developed a sensor-agnostic framework leveraging pre-trained machine learning models. We also generated land cover models for Sentinel-2 (20m) and Landsat 8 imagery (30m) using either a single date of imagery or two dates of imagery for mapping land cover type. The two-date model includes 11 land cover type classes, whereas the single-date model contains 6 classes. The models' overall accuracies were 84% (Sentinel-2 single date), 82% (Sentinel-2 two date), and 86% (Landsat 8 two date) across the continental United States. The three different models were built into an ArcGIS Pro Python toolbox to enable a semi-automated workflow for end users to generate their own land cover type maps on demand. The toolboxes were built using parallel processing and image-splitting techniques to enable faster computation and for use on less-powerful machines.

15. SUBJECT TERMS

Land cover--Remote sensing, Machine learning, Python (Computer program language), ArcGIS, Geographic information systems

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	None	47	