Graduate Theses, Dissertations, and Problem Reports

2021

Assessing Synthetic Aperture Radar (SAR)-Derived Temporal Patterns and Digital Terrain Data for Palustrine Wetland Mapping

Jaimee L. Pyron
jp0160@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Remote Sensing Commons

Assessing Synthetic Aperture Radar (SAR)-Derived Temporal Patterns and Digital Terrain Data
for Palustrine Wetland Mapping

Jaimee Pyron

Thesis submitted to Eberly College of Arts and Sciences at West Virginia University for partial
fulfillment of the requirements for the degree of Master of Arts in the Department of Geology
and Geography

Aaron Maxwell, PhD., Chair

Brenden McNeil, PhD

Michael Strager, PhD

Department of Geology and Geography

Morgantown, West Virginia

2021

ABSTRACT

Assessing Synthetic Aperture Radar (SAR)-Derived Temporal Patterns and Digital Terrain Data
for Palustrine Wetland Mapping

Jaimee Pyron

Palustrine wetland systems are important ecosystems and provide numerous ecosystems services to support society. Unfortunately, they remain under constant threat of devastation due to land use practices and global climate change, which underscores the need to identify, map, and monitor these landscape features. This study explores harmonic coefficients and seasonal median values derived from Sentinel-1 synthetic aperture radar (SAR) data, as well as digital elevation model (DEM)-derived terrain variables, to predict palustrine wetland locations in the Vermont counties of Bennington, Chittenden, and Essex. Support vector machine (SVM) and random forest (RF) machine learning models were used with various combinations of the three datasets: terrain, SAR seasonal medians, and SAR harmonic time series coefficients. For Bennington County, using the harmonic and terrain data with a RF model yielded the most accurate results, with an overall accuracy of 76%. The terrain data alone and RF model produced the highest overall accuracy in Chittenden County with an accuracy of 85%. In Essex County any combination of the three datasets and the RF model yielded the highest overall accuracy of 81%. Generally, this study documented better performance using the RF algorithm in comparison to SVM. Terrain variables were generally important for differentiating wetlands from uplands and waterbodies. However, Sentinel-1 data, represented as harmonic regression coefficients and seasonal medians, provided limited predictive power. Although Sentinel-1 SAR data were of limited value in the explored case studies, findings may not extrapolate to other SAR datasets using different polarizations, wavelengths, and/or spatial resolutions.

Contents

1. Introduction

Wetlands are a prime example of the coupling of environmental and anthropogenic systems and how they can serve society. They are abundant across urban and rural landscapes, where they provide ecosystem services such as carbon sequestration, wildlife habitation, water retention from flooding, and recreation [1–3]. Of these wetland ecosystems, palustrine wetlands serve as a necessary part of the ecosystem. Palustrine wetlands are inland and vegetated, and include but are not limited to bogs, fens, marshes, or swamps and are dynamic ecosystems that bear economic and ecological importance [4–6].

While these ecosystems are rich in resources, they are no longer rich in abundance when compared to how prevalent they have been historically—wetlands are disappearing at an astonishing rate as a direct result of land use practices associated with urbanization and agriculture alongside global climate change issues such as deforestation and sea level rise [2,7]. For this reason, they have become a recent topic of discussion and study across both the environmental and social sciences. At the same time, technologies have developed and evolved, offering advancing opportunities to map and monitor these landscape features. Nevertheless, identifying wetlands using remote sensing and surveying techniques have proven difficult as they often lack a single unifying feature (i.e., a characteristic spectral signature) and definitive boundaries [8,9]. Essentially, wetlands often have inherently fuzzy or gradational boundaries with surrounding uplands, resulting in inherent errors in "hard" boundaries and delineations. Further, the highly dynamic nature of wetlands (i.e., variety of site characteristics and seasonal variability in vegetation and moisture conditions) and technological limitations of satellite-based Earth observation has made mapping them an arduous task. Just a few of the potential limitations include data cost, sub-optimal weather conditions, image resolution, and satellite launch date. While wetland delineation datasets are maintained by the state and federal government, flaws exist in these datasets that emphasize the need for automated identification methods with increased accuracy and frequent updates—especially as climate and land use change continue [8,10,11].

While there has been an increase in public interest and scientific research relating to wetlands in recent years, research has mainly focused on coastal environments, thus creating a need for further study of inland wetlands [8]. This is underscored in a study by Davidson [2] in

1

which an analysis of 20th century data showed an estimated 69-75% loss of inland wetlands compared to a 46-50% loss of coastal wetlands due to human development, including a rise in global population, increased and intensive agriculture practices, and urban encroachment on natural lands [2]. This exemplifies the need for multi-temporal wetland studies, rather than evaluating a single date. Further, studying wet-land presence in a dense time series substantiates trends of long-term wetland loss while making models more robust by adding a temporal element, or potentially improving mapping accuracy by characterizing seasonal signatures and variabilities. Due to the seasonality of wetlands, such as annual changes in moisture levels and the presence of standing water, as well as their historic susceptibility to climate change, it is pertinent to consider time series analysis during wetland identification. Moreover, a long-term time series analysis allows climate scientists to forecast potential wetland variability, therefore helping policymakers enact regulatory measures to preserve these areas.

To understand the importance of wetlands, the public must first understand to what extent they are disappearing in the United States. A national-level wetland inventory exists but has known flaws and gaps. The National Wetlands Inventory, or NWI, is a publicly available dataset provided by the US Fish and Wildlife Service (USFWS). These wetland areas are delineated primarily using interpretation of aerial orthophotography based on a classification system developed by Cowardin et al. [12] over 40 years ago. Cowardin et al. [12] differentiates five wetland types at the highest classification hierarchy (systems): marine, estuarine, riverine, lacustrine, and palustrine. These systems are further subdivided into subsystems based on hydrological, chemical, biological, and geomorphologic factors [12]. Palustrine wetlands specifically include non-tidal wetlands and tidal wetlands with a salinity less than 0.5% that are often characterized by their vegetation, including scrub/shrub, forests, emergent mosses, emergent lichen, and persistent emergent [12]. These areas are often bounded by other wetland systems or up-lands [12].

Previously mapped wetlands in the NWI dataset are not regularly updated, meaning they may not be representative of the wetland's current size and spatial extent. Aside from these administrative errors, the NWI suffers from errors within the dataset itself. It is known to have a low commission but high omission error, meaning that while it does not often incorrectly include non-wetlands in the dataset (i.e., commission error), existing wetlands are missed (i.e., omission

error). Sharpe et al. [13] also documented that the NWI generally has low levels of both classification commission and omission of wetlands greater than 0.40 hectares, implying that it does not accurately portray smaller wetlands. Tiner [14] reiterates this point by addressing the issue of the target mapping unit, or tmu. The tmu is the smallest wetland consistently mapped, although not necessarily the smallest wetland present [14]. This highlights a need to further improve classifications, as smaller palustrine wetlands hold economic and environmental importance but are consistently and knowingly overlooked and unmapped. Maintaining an accurate and up-to-date inventory of wetlands is critical in a time of global climate and land use change. For example, Wilkins et al. [15] noted that an awareness of wetlands positively related to concern for their loss.

Evaluating wetlands using remote sensing techniques often includes the use of optical, synthetic aperture radar (SAR), and/or terrain data. Traditionally, remote sensing techniques have relied on optical imagery with bands in the visible, near infrared (NIR), and shortwave infrared (SWIR) spectral ranges for wetland classification analysis, although such data have limited availability due to low temporal resolution and cloud contamination [8]. Studies using optical imagery for wetland identification have recorded accuracies above 80% from WorldView-2 commercial imagery, although this data source may not be viable for large-area studies or those with a budgetary constraint [16]. Amani et al. [17] obtained a lower overall accuracy of 71% using Landsat-8 Operational Land Imager (OLI) imagery, a widely used free alternative. This low accuracy highlights the unsuitability of only optical imagery for wetland classification and the need for an open-source solution for future wetland mapping. While the normalized difference water index (NDWI) suggests that the green and near infrared bands are best for identifying waterbodies, this may not extrapolate to wetlands in general. Corcoran et al. [18] found that the blue and red bands were the most important input bands for classifying wetlands. Also, standing water is not consistently present in palustrine wetlands; instead, the spectral response is more associated with vegetation that has adapted to moist soil conditions with seasonal or sporadic inundation. This variability, coupled with the temporal and atmospheric limitations of optical imagery, supports a need for further research focused on wetland mapping using SAR data.

While research suggests that optical imagery alone is insufficient for wetland identification and delineation, evidence suggests that the combined use of alternate or multiple datasets enhances results [18–20]. Sentinel-1 SAR imagery is an open-source alternative that can be used regardless of weather conditions. SAR is a type of active satellite technology that creates a finer spatial resolution image by mimicking the motion of a longer antenna, hence synthetic aperture [21]. SAR sensors operate in the microwave portion of the electromagnetic spectrum and use longer wavelengths than passive, multispectral sensors. These wavelengths are cloud penetrating, allowing for repeat data collection on a regular basis, regardless of atmospheric conditions [21]. Aside from this, SAR is unique from traditional satellite imagery in that it is side-looking, meaning that the microwave pulses are emitted and received at an angle in order to measure the backscatters time of travel back at the sensor [21]. C-band SAR, which is implemented in the Sentinel-1, is advantageous when identifying wetlands under suboptimal atmospheric conditions, as it operates at a wavelength range between 7.5 and 3.75 centimeters, making it cloud-penetrable [21,22].

Polarization states are determined by the way in which the microwave signals are transmitted and received, and there is debate surrounding the optimal polarization for wetland identification. Recent studies have recognized HH-polarization as ideal for wetland mapping, but it is only available for polar and subpolar regions [17,23]. Baghdadi et al. [24] suggest that cross polarization has been identified as providing the best separation between individual wetlands, which is more widely available than HH-polarization. SAR has been noted as being particularly apt for identifying waterbodies, as Corcoran et al. [18] document in a multi-sensor, multi-temporal wetland study that SAR specifically improved accuracy when differentiating between wetlands and uplands. This can be largely attributed to SAR being situated in the microwave region of the electromagnetic spectrum, and subsequently the influence of terrain roughness and the dielectric constant on the microwave backscatter [21,25]. Rough surfaces generate higher backscatter due to a phenomenon known as volume backscattering, which ultimately sends more backscatter to the sensor, resulting in a strong signal [21]. Still waterbodies lack surface roughness and result in specular reflectance, where little energy is scattered back toward the sensor, resulting in a lower backscatter coefficient [21,25]. Essentially, backscatter coefficients are dependent on hydrological processes and tend to exhibit a strong correlation to seasonality due to changes in specular reflectance associated with varying inundation extents and soil

moisture levels [26]. Issues with SAR wetland classifications arise with the presence of vegetation, as vegetated waterbodies suffer most from the issue of double bounce [22,27]. This differs from the typical specular reflection seen in most waterbodies, which returns a signal after a single bounce, in that the incoming radar pulses make contact with several objects before returning a signal [22,27]. While SAR has disadvantages, such as a reputation as being difficult to interpret and signals that are made complex by the presence of vegetation in wetlands, using complimentary data and certain modelling techniques have made it a viable option for identifying wetlands.

Terrain variables, often derived from digital elevation models, are also used in wetland classification. Maxwell et al. [10] used terrain-derived variables such as slope, surface curvature, and profile curvature, to detect wetland probability using a random forest (RF) machine learning classifier [10]. Notably, this study found that terrain variable importance differs based on geographic region [10]. Coupling these terrain-derived variables with SAR and optical imagery in multi-sensor and multi-frequency studies can yield robust results wherein the respective disadvantages of each dataset are partially offset by the complimentary data [18,28]. Banks et al. [29] used RADARSAT-2, DEM, and DSM data to classify shallow water, marsh, swamp, water, forest, and non-forested areas. By fusing SAR, DEM, and DSM data, the researchers were able to achieve user's and producer's accuracies in excess of 90% for each land cover class [29]. One notable limitation of terrain data is the temporal aspect; however, this is not typically an issue as topography is slow to change, but it could be relevant for quickly changing hydrologic features, such as temporary wetlands.

The use of time series data derived from multi-data collections in wetland remote sensing research can allow characterization of seasonal patterns and changes in moisture levels and vegetation. The Landsat data archive is well suited for time series analysis, as the first satellite launched in 1972, allowing for over 40 years of landscape change analysis. However, such lengthy archives are rare, again hindering studies that use solely optical data [30]. SAR and high-resolution digital elevation models have only become freely available in recent years (i.e., over the last decade), limiting the ability to use these datasets in dense time series analyses to characterize non-static relationships and landscape change. However, a shorter time series

covering only a few consecutive years is generally adequate for investigating seasonal patterns, which is of interest here.

Seasonal variations can be captured using SAR data, as cloud penetration allows for uninterrupted data collection at a defined and consistent return interval. Schlaffer et al. [26] demonstrated the use of parameters derived from a harmonic time series of SAR backscatter coefficients over two hydrological years [26]. They were able to extract seasonal signatures and use unsupervised classification to identify wetlands in Zambia [26]. In the case of intermittent wetlands, there may only be a few days or weeks when site conditions allow them to be differentiated from surrounding uplands, in which case having more frequent image collection dates increases the chances of detection [31,32].

The objectives of this study are to (1) assess measures derived from a time series of Sentinel-1 C-band SAR for mapping palustrine wetlands and differentiating them from uplands and water bodies, (2) compare harmonic regression and seasonal aggregating techniques for summarizing a SAR time series for palustrine wetland mapping, and (3) assess DEM-derived variables for palustrine wetland mapping. This study makes use of wetland data available for the state of Vermont. Palustrine wetland mapping across three counties in the state were investigated as separate case studies.

2. Materials and Methods

2.1 Study Area

The study area is comprised of nine major biophysical regions characterized by valley, highland, and piedmont regions (Figure 1). This research was formatted as a case study within three counties in the state of Vermont, USA: Chittenden, Bennington, and Essex. Vermont is located in the northeastern United States and borders Canada to the north, New Hampshire to the east, Massachusetts to the south, and New York to the west. Vermont is primarily forested with numerous lakes throughout the state resulting from recent glacial activity [33,34]. These counties were chosen based on a combination of diverse physical and social influencing factors. The state has 14 counties and a total population estimated at 624,000 [35]. Chittenden County is the most populated county in Vermont with an estimated population of 164,000 and the county seat,

Burlington, is notably the most populated city in the state [35]. Chittenden lies in the northwestern part of the state bordering Lake Champlain to the west. Bennington County has an estimated population of 35,000 and lies in the southernmost part of the state, bordering New York and Massachusetts [35]. Green Mountain National Forest encompasses most of Bennington, lending to its mountainous terrain. Essex County has an estimated population of 6,000, the lowest of any county in Vermont and the entirety of New England. Essex County was chosen as it is a rural county with both a small population and low population density, which provides a contrast to Chittenden and Bennington. Notably, Vermont was chosen because of the quality of the state-made VSWI dataset as well as the thorough wetland regulations within the state.
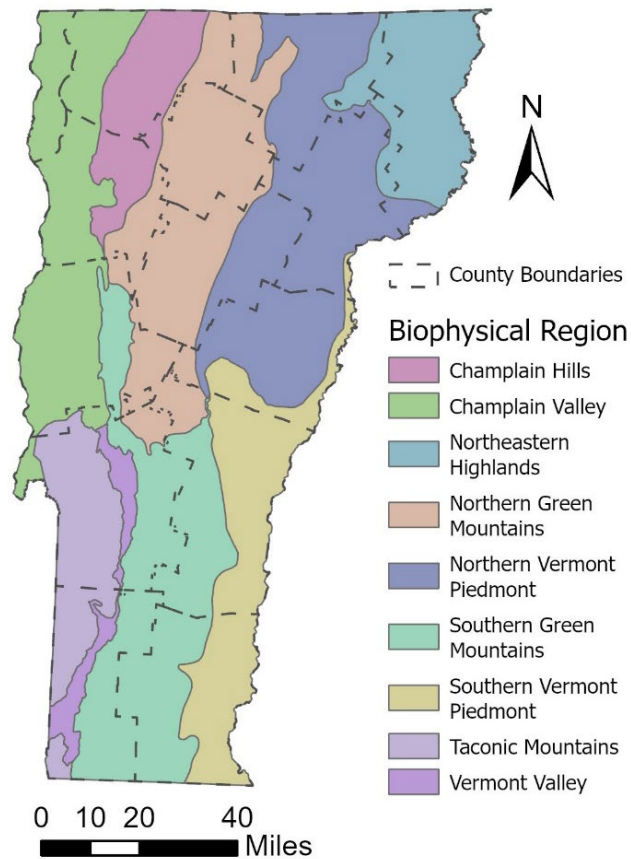
2.2 Data

This study used Sentinel-1 SAR imagery and the Vermont Significant Wetlands Inventory, or VSWI, dataset. Sentinel-1 SAR is comprised of a constellation of two satellites (1A and 1B) that were launched on April 3, 2014 and April 25, 2016, respectively [25]. The Sentinel-1 specifications used for this study are provided in Table 1. SAR GRD data were collected and preprocessed approximately every 6 days for the full calendar years of 2017 through 2019 using Google Earth Engine (GEE). Notably, GRD is a multi-look product that has been focused to ground range, as opposed to the traditional single-look, slanted format [21]. Ground range measurements are a correction of the original slanted coordinates recorded by the satellite that have been projected onto an ellipsoidal model of the Earth, which allows the data to be visualized in a projected map space [21,25]. While this conversion results in the loss of phase information and reduced spatial resolution, it provides square pixels and reduced radar speckle [21]. The Sentinel-1 products used were collected in VV and VH dual polarized modes, meaning that the former is vertically transmitted and vertically received, wherein the latter is vertically transmitted and horizontally received. Specifically, data were collected in Sentinel-1's Interferometric Wide mode at 25-meter spatial resolution, where each pixel value represents a backscatter coefficient for each unit of ground area in decibels. The decibel is a normalized value that represents the observed strength, or amplitude, of the backscatter signal [21,25]. This reading indicates whether the incident microwaves are being reflected toward the sensor or away from it. Preliminary data collection indicated that VH-polarized and VV-polarized images were collected on 159 dates in Bennington County, 163 dates in Chittenden County, and 166 dates in Essex County throughout the 2017 to 2019 calendar years. An example of these SAR images in Essex County is shown in Figure 2.

Specification	Value
Mode	Interferometric Wide
Product	GRD
Wavelength	5.6 cm
Frequency (GHz)	5.405
Resolution	25 m
Polarizations	VV, VH
Incidence Angle Range	29.1° - 46.0°
Revisit Time Range	6-12 days

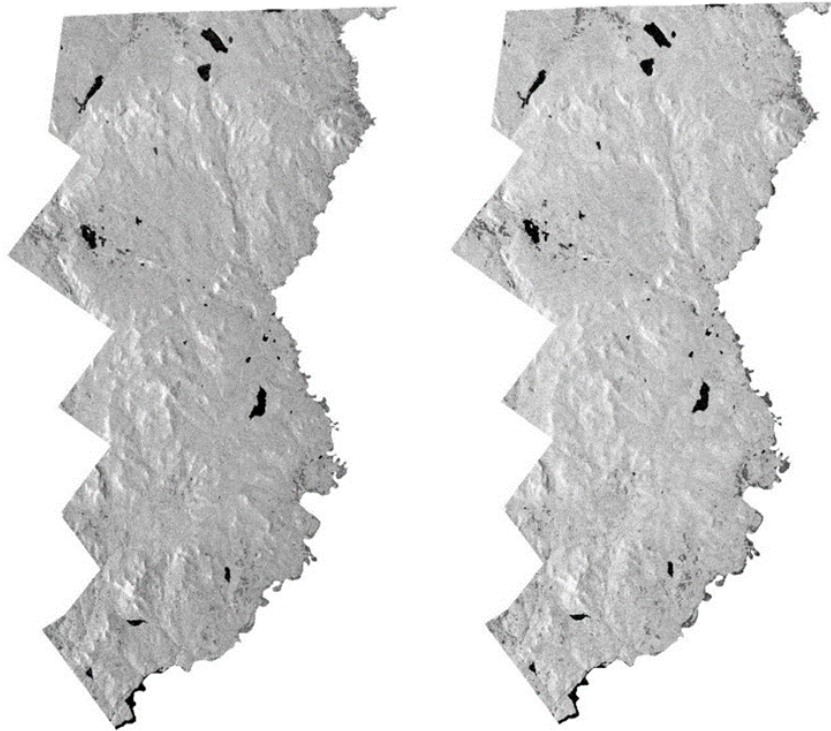Table 1. Sentinel-1 SAR specifications [21]

Figure 2. Example Sentinel-1 SAR data for Essex County. (a) VV Polarization and (b) VH Polarization. Lighter shades indicate higher backscatter.

County contains 266 acres of Class I wetlands and 22,123 acres of Class II wetlands (Figures 3 and 4) [38]. While this does not directly affect the analysis, it helps gauge the abundance of wetlands in the study areas. Furthermore, the dataset includes a few lacustrine and riverine wetlands that were manually examined and removed since the study focuses on palustrine wetlands. This was accomplished using manual image interpretation of multiple Landsat 8 Operational Land Imager (OLI) images using the dates shown in Table 2. Multiple Landsat images are needed to cover the entire study area and account for potential cloud cover.

Bennington	Chittenden	Essex
October 2, 2017	October 2, 2017	October 4, 2017
October 27, 2017	October 27, 2017	October 27, 2017
November 28, 2017	November 28, 2017	November 11, 2017

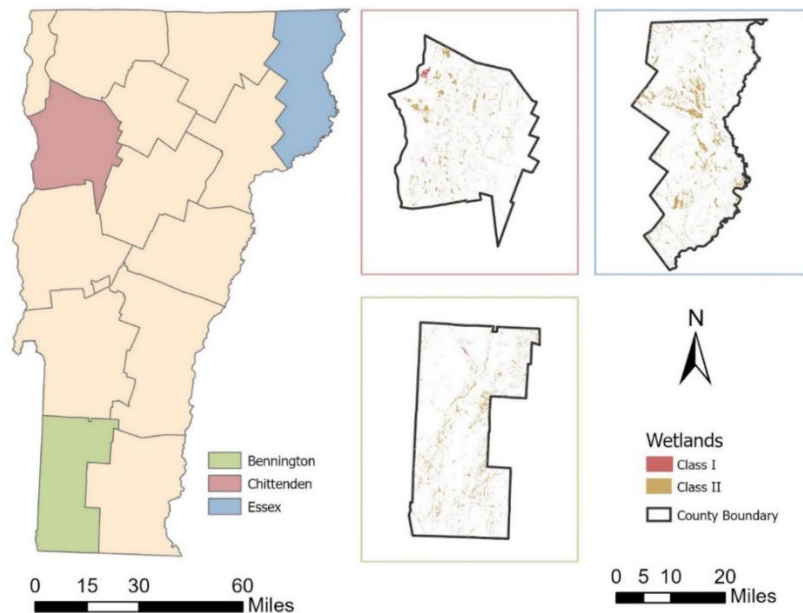Table 2. Dates of Landsat image collection for manual interpretation.

Figure 2. Locations of Bennington, Chittenden, and Essex counties and their Class I and Class II wetland coverage.
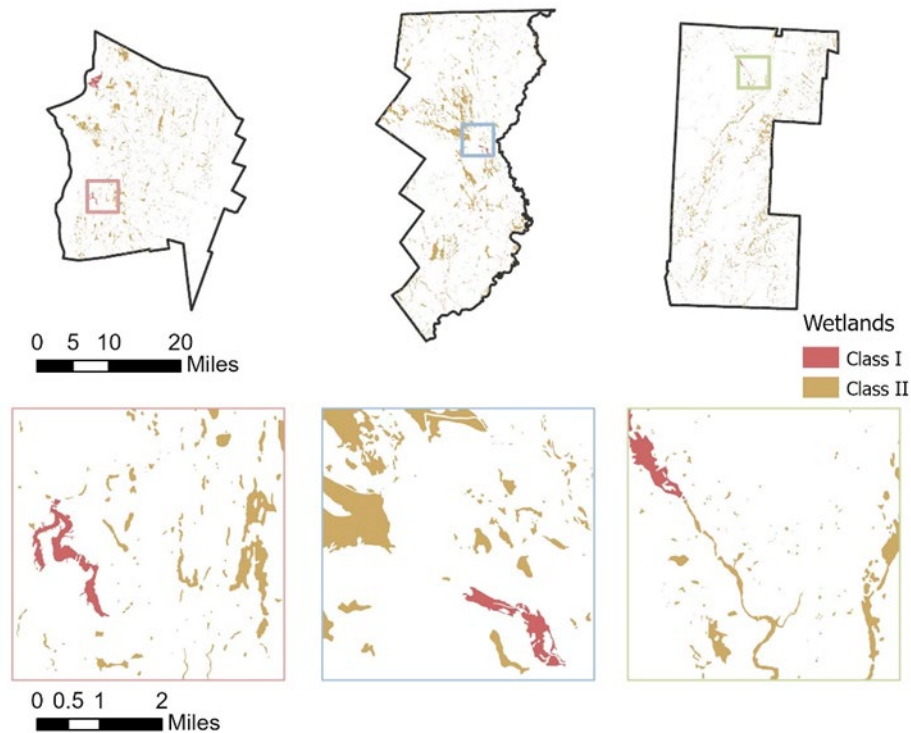
Figure 3. Example Class I and Class II wetlands in the three counties of interest.

2.3 Exploratory Data Analysis

Prior to conducting the classification, the data were subject to exploratory data analysis. This included simple data visualization using boxplots, as well as implementing the Kruskal-Wallis test. The Kruskal-Wallis test is a non-parametric test that relies on ranks and assesses whether there are differences between three or more classes in regard to the values of a continuous variable and is a non-parametric alternative to analyses of variance (ANOVA). For this study, Kruskal-Wallis was used to test the differences in predictor variable values between the mapped groups, i.e. upland, wetland, and waterbodies. This requires using p-values, and although not all encompassing, they are a useful tool for statistical analyses [39].

2,4 Harmonic Time Series Analysis

Due to the seasonal nature of this study, a non-linear harmonic time series analysis was implemented. Two sets of sampled pixel backscatter coefficients, one for VV polarization and one for VH polarization, were used as the predictor variables. At a conceptual level, seasonal oscillations of a time series can be described using Equation 1, and simplified to Equation 2, below:

$$y = m + a\cos(\omega t) + b\sin(\omega t), \quad (1)$$

$$y = m + c\cos(\omega t - \varphi), \quad (2)$$

In these equations m represents the mean of the series, c represents amplitude, φ represents phase angle, t represents time, and ω represents the oscillation frequency [40]. Harmonic time series are sinusoidal in nature and the appropriate number of wavelengths fitted to the model can be determined through trial-and-error while considering the implications on the resulting temporal resolution [40]. For this study, the amplitude and phase angle for VV backscatter and VH backscatter were estimated, resulting in four total variables for use in the classification models. This model is ideal for SAR data, as radar is cloud-penetrable and sensitive to changes in moisture content, therefore lending to its suitability for seasonal analysis.

The time series allows for visualization, characterization, or quantification of the effects of seasonality. With well over 100 images collected over a 3-year time frame in each county, this study can be characterized as making use of a dense time series. An example is shown in Figure 4 of the VV and VH backscatters in Essex county. Dense time series analyses are a relatively new technique for analyzing SAR data, as they require a large number of data points collected over time and have only been made more possible in recent years due to the Sentinel-1 mission's frequent return interval, once every 6-12 days [26]. The open-source R programming language and software environment can perform time series analysis using both its baseline functions and the forecast package [41,42].
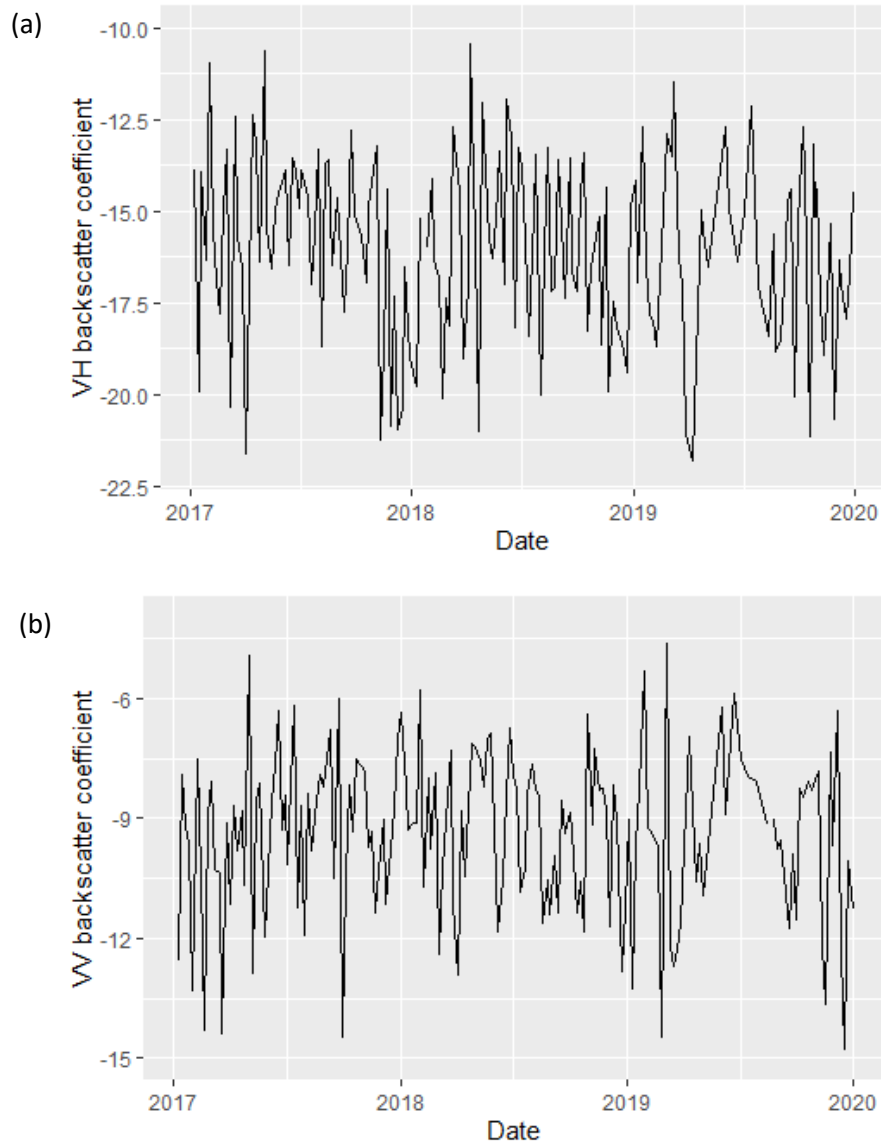
(a)

(b)

Figure 4. Example palustrine wetland time series for (a) VH and (b) VV polarization

The terrain variables were derived from the Shuttle Radar Topography Mission (SRTM) digital elevation models (DEMs) at 30-meter spatial resolution using ArcGIS Pro and the ArcGIS Geomorphometry and Gradient Metrics toolbox [43]. A total of six terrain variables, shown in Table 3, were included in this study: slope, slope position, roughness, plan curvature, profile curvature, and topographic wetness index. Slope is the gradient of topographic change or the first derivative of the elevation surface. Slope position is calculated by subtracting the mean

of the elevation values within a moving window from the center cell elevation value. High values indicate higher topographic positions, such as ridges, while low values indicate lower positions, such as valleys [44,45]. Roughness is estimated as the variance in elevation measurements within a local window. [46]. Plan and profile curvatures refer to the curvature perpendicular to and in the direction of maximum slope, respectively [47,48]. Compound topographic wetness index is a measure of steady state moisture that takes into account the amount of contributing area to each cell, as approximated with flow accumulation, and the topographic slope at the cell. Cells with a larger contributing area and a shallower slope would be expected to have a larger moisture content in comparison to cells with a smaller contributing area and/or steeper slope [49]. For slope position and roughness, a circular window with a radius of 9 cells was used.

Variable	Description
Slope (°)	Surface gradient, measured in degrees
Slope Position	Second derivative of slope
Roughness	Elevation difference between adjacent DEM cells
Plan Curvature	Curvature perpendicular to slope
Profile Curvature	Curvature in direction of slope
Compound Topographic Wetness Index (CTWI)	Steady state, topographically grounded wetness index

Table 3. DEM-derived terrain variables

2.5 Machine Learning Models

Wetland classification was conducted with two machine learning techniques: support vector machine (SVM) and random forest (RF). Support vector machine is a binary classifier that focuses on the identification of boundaries between two classes using a hyperplane and the training samples, or support vectors, that are used to define it. In order to model nonlinear boundaries, the data can be projected into a higher dimensional feature space where boundaries

may be more linear separable. This is known as the "kernel trick". Also, methods have been developed to allow for multi-class classification and not just binary separation [50].

Random forest is an ensemble decision tree classifier that uses a "majority-wins" system on many individual decision trees [51]. Each tree in the ensemble uses a subset of the available training samples, selected using bagging (i.e., random sampling with re-placement) as opposed to the entire set. Also, only a random subset of predictor variables is available for splitting at each decision node. The goal of using a subset of the training samples in each tree and a random subset of training samples at each node is to reduce the correlation between trees in the model, which can result in improved model generalization due to reduced correlation between the trees [52,53]. SVM and RF are both non-parametric, meaning they do not make strong assumptions about the data, most importantly regarding its normality. Further, they have become standard methods for classification in remote sensing [53].

Each machine learning technique has the same inputs, tested over 6 separate experiments, shown in Table 4. Differentiated categories include palustrine wetlands, uplands, and remaining waterbodies. This categorization is common in studies that attempt to identify palustrine wetlands using machine learning, as the backscatter coefficient of uplands and the remaining wetlands can be markedly different [18]. However, further segmentation could lead to confusion among similar classes [17]. Random samples were collected for each classification type, with 75% of the data used for training and 25% of the data used for validation. With 600 data points for each county, or 200 samples per class, this resulted in 450 samples used for training and 150 used for model validation. The boundaries from the VSWI wetlands dataset were used to delineate the palustrine wetlands, the VSWI water classes dataset was used to delineate other waterbodies, and all excluded areas were used to sample uplands.

In order to optimize algorithm hyperparameters, I used 5-fold cross validation in which the data were split into five, non-overlapping partitions. Models were then trained using four folds and withholding the remaining fold for validation. Training is implemented five times, each time holding out one of the five folds. The hyperparameters providing the best performances are determined based on the best average performance, as measured using Kappa, on the withheld data.

Once models were optimized, the best tested hyperparameters were then used to train the final model. This final model was then used to predict the withheld validation data. Specifically, individual classification accuracies were assessed through a confusion matrix and measures of overall accuracy and Kappa. Comparative classification performance was assessed using McNemar's test, a non-parametric test for nominal data. All machine learning analyses, hyperparameter tuning, and model assessment were performed using the R open-source software and the caret package [41,54].

	SAR Harmonic	SAR Median	Terrain-Derived
Model 1	X		
Model 2		X	
Model 3			X
Model 4		X	X
Model 5	X		X
Model 6	X	X	X

Table 4. Datasets used for each model

3. Results

3.1 Bennington County

The exploratory data analysis produced insight into the individual variable distribution and importance of using boxplots and the Kruskal-Wallis test. The distribution of the VV and VH harmonic coefficients are shown in Figure 5. The boxplots show numerous outliers amongst the VV amplitude and VH amplitude values, whereas the VV Phase and VH Phase boxplots visually show less variation between classes. This is illustrated in Table 4, as both VV amplitude and VH amplitude have a statistically significant p-value of less than 0.05, whereas VV phase and VH phase do not reach the threshold for statistical significance.
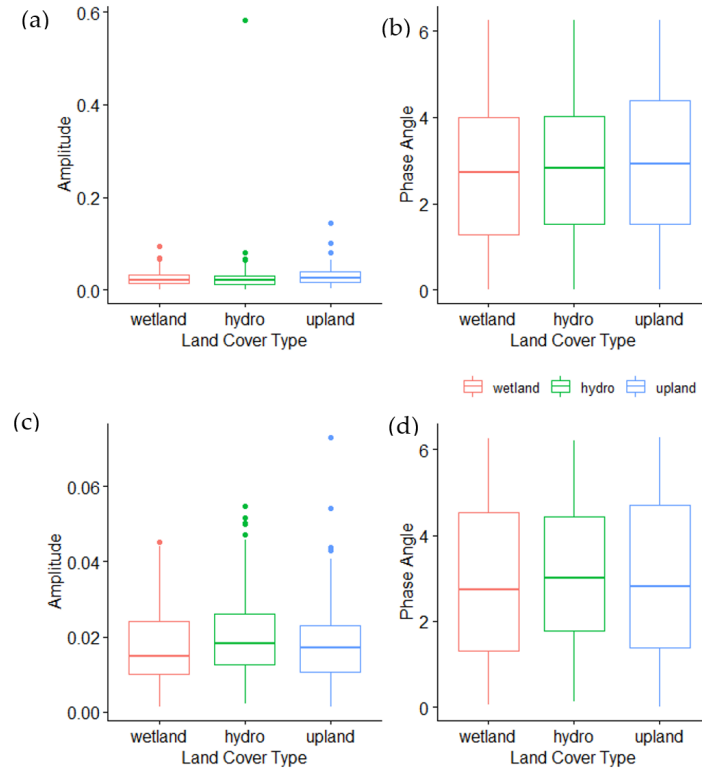
Figure 5. Boxplots of distribution for (a) VV Amplitude, (b) VV Phase, (c) VH Amplitude, and (d) VH Phase by mapped class.

Variable	p-value
VV Amplitude	0.00
VV Phase	0.51
VH Amplitude	0.02
VH Phase	0.36

Table 4. Resulting p-values from Kruskal-Wallis test for VV Amplitude, VV Phase, VH Amplitude, and VH Phase.

Exploratory data analysis was also performed on the terrain variables of profile curvature, slope, plan curvature, slope position, roughness, and TWI. The boxplot distributions are shown in Figure 6 and the resulting p-values from the Kruskal-Wallis test are shown in Table 5.

Visually, outliers are present in each of the land cover classes across all terrain variables. However, the p-values for all terrain variables except plan curvature are still below 0.05, therefore meeting the threshold for statistical significance.
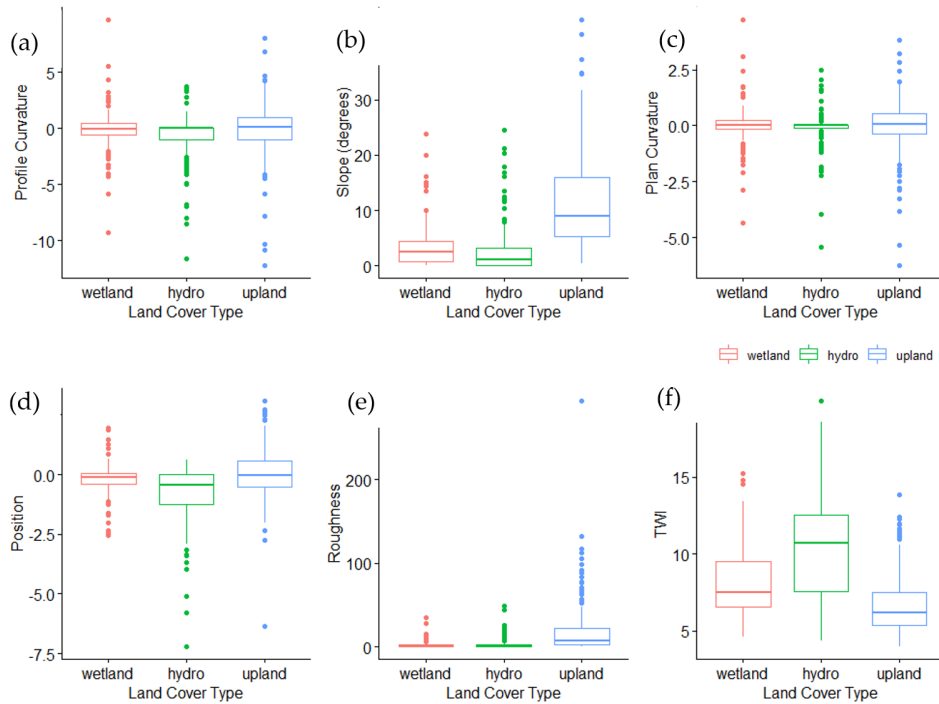

Figure 6. Boxplots of distribution for (a) Profile Curvature, (b) Slope, (c) Plan Curvature, (d) Slope Position, (e) Roughness, and (f) TWI by mapped class.

Variable	*p*-value
Profile Curvature	0.00
Slope	0.00
Plan Curvature	0.06
Position	0.00
Roughness	0.00
TWI	0.00

Table 5. Resulting p-values from Kruskal-Wallis test for terrain variables.

Finally, the random forest and support vector machine models were run for the six model combinations: harmonic only, median only, terrain only, harmonic and terrain, median and terrain, and all variables. The overall accuracies and Kappa statistics for those models are shown in Table 6. Regarding the random forest model, the harmonic and terrain model had the highest accuracies with an overall accuracy and Kappa statistics of 0.76 and 0.64, respectively. Of the models that only took one type of variable, the terrain variables outperformed the others with an overall accuracy of 0.75 and a Kappa statistic of 0.63. The harmonic model had the poorest performance with an overall accuracy of 0.37 and a Kappa statistic of 0.06. In addition to traditional accuracy metrics, the F1 score, precision, and recall for the palustrine wetland class are shown in Figure 7. It is worth noting that the hydro and upland classes consistently had a higher F1 score than the wetland class. The SVM models had slightly lower accuracies across the board, except for any model where the median values were involved. The harmonic and terrain model as well as the median and terrain model had the highest SVM accuracies, with an overall accuracy of 0.73 and a Kappa statistic of 0.59. Of the models that only used one group of variables, the terrain only model once again outperformed the others with an overall accuracy of 0.71 and a Kappa statistic of 0.57. Figure 8 shows F1 score, precision, and recall for the wetland class and SVM models.

Model	RF OA	RF Kappa	SVM OA	SVM Kappa
Harmonic only	0.37	0.06	0.35	0.02
Medians only	0.55	0.33	0.59	0.38
Terrain only	0.75	0.63	0.71	0.57
Harmonic and Terrain	0.76	0.64	0.73	0.59
Median and Terrain	0.73	0.60	0.73	0.59
All Variables	0.74	0.61	0.72	0.58

Table 6. Overall accuracy (OA) and Kappa statistics for random forest (RF) and support vector machine (SVM) classifications.
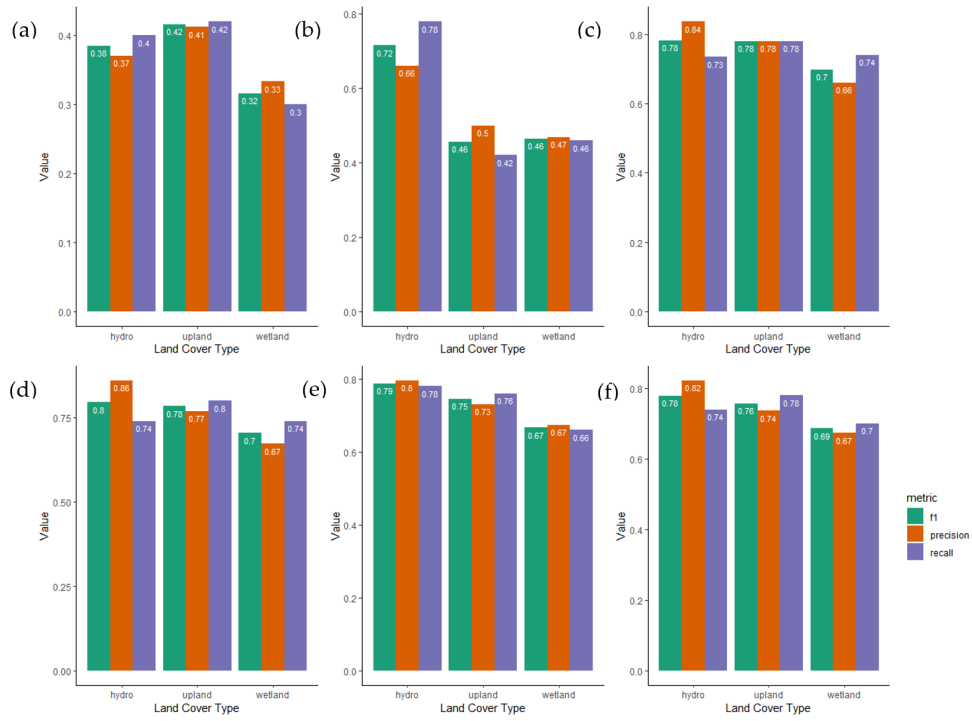
Figure 7. Random forest F1 score, precision, and recall by class when using (a) SAR harmonic variables, (b) SAR medians, (c) terrain variables, (d) harmonic and terrain variables, (e) median and terrain variables, and (f) all variables.
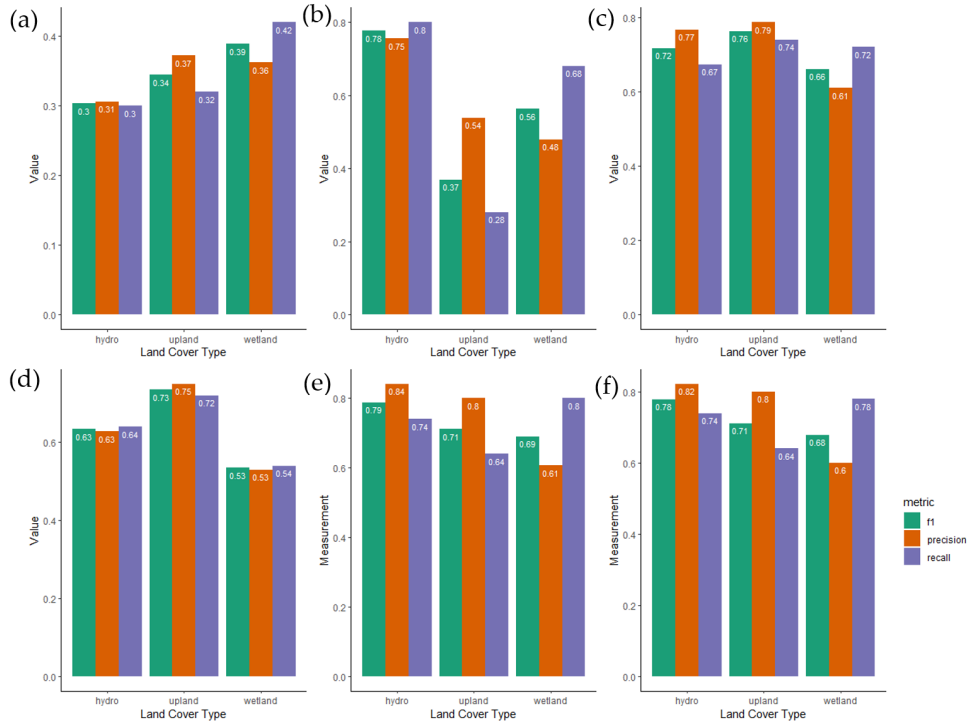
Figure 8. Support vector machine F1 score, precision, and recall by class when using (a) SAR har-monic variables, (b) SAR medians, (c) terrain variables, (d) harmonic and terrain variables, (e) median and terrain variables, and (f) all variables

3.2 Chittenden County

The distribution of the VV and VH harmonic coefficients for Chittenden county are shown in Figure 9. Visually, outliers can be seen in all land cover types in the VV amplitude dataset and for the hydrological class for the VH amplitude dataset. This set also has less variation between the classes, as visually interpreted in the boxplots and shown by the p-values in Table 8. None of the harmonic variables reach statistical significance when analyzed in Chittenden county, with VH amplitude notably being just above the 0.05 threshold at a value of 0.06. This is later reflected in model performance, but these values were still modeled to show a comparison across all three counties.
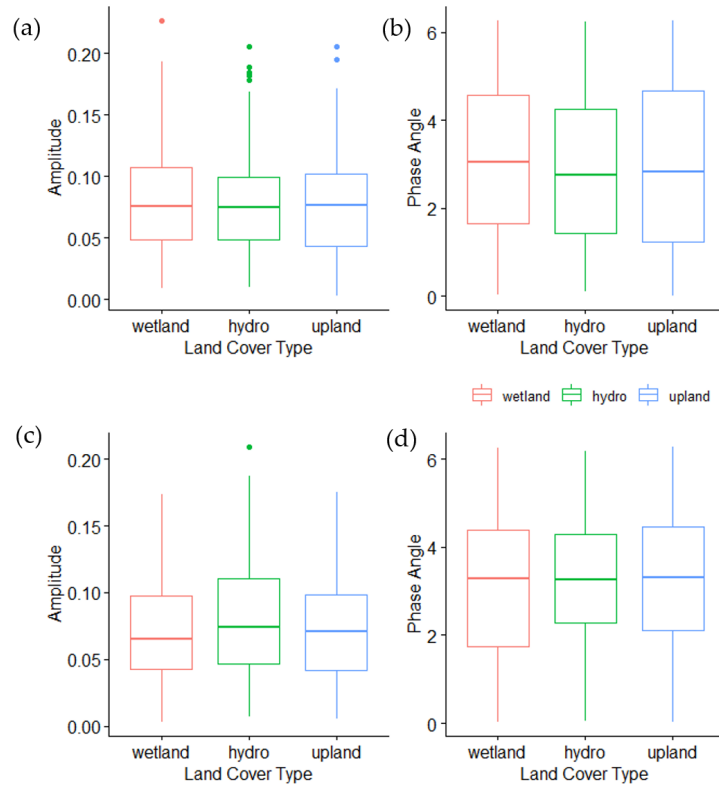
Figure 9. Boxplots of distribution for (a) VV Amplitude, (b) VV Phase, (c) VH Amplitude, and (d) VH Phase by mapped class.

Variable	*p*-value
VV Amplitude	0.62
VV Phase	0.47
VH Amplitude	0.06
VH Phase	0.85

Table 7. Resulting p-values from Kruskal-Wallis test for harmonic variables.

Like Bennington county, the boxplots for Chittenden are shown in Figure 10 with outliers for each land cover type in each model. The Kruskal-Wallis p-values for these varied, as shown in Table 8. Profile curvature and plan curvature were not statistically significant, with p-values of

0.34 and 0.87, respectively. The other four variables had low enough values to reach a level of statistical significance. While plan curvature had not reached the threshold for statistical significance previously in Bennington county, profile curvature had. This is worth noting as one of the differences in the performance of these measures across different geographical regions.
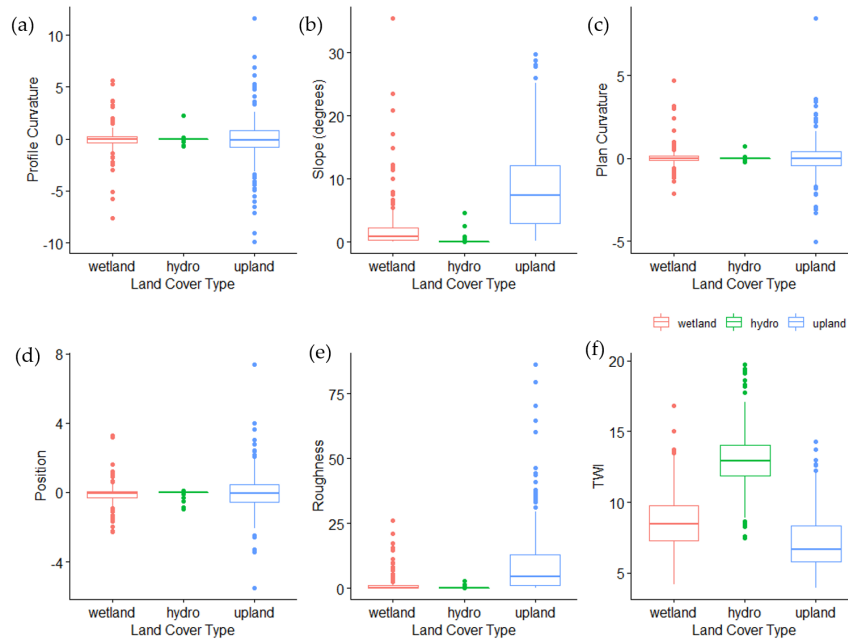

Figure 10. Boxplots of distribution for (a) Profile Curvature, (b) Slope, (c) Plan Curvature, (d) Position, (e) Roughness, and (f) TWI by mapped class.

Variable	*p*-value
Profile Curvature	0.34
Slope	0.00
Plan Curvature	0.87
Position	0.02
Roughness	0.00
TWI	0.00

Table 8. Resulting p-values from Kruskal-Wallis test for terrain variables.

The random forest and support vector machine models were run for the six model combinations. The overall accuracies and Kappa statistics for those models are shown in Table 9. Referencing the random forest model, the terrain variables outperformed all of the other models with an overall accuracy of 0.85 and a Kappa statistic of 0.78. The harmonic model performed the worst with an overall accuracy of 0.33 and a Kappa statistic of 0.01. In addition to traditional accuracy metrics, the F1 score, precision, and recall for the wetland class are shown in Figure 11. The SVM models had slightly lower accuracies across the board, except with the harmonic only model. The model that included all variables and the terrain only model had the highest of the SVM accuracies, with overall accuracies of 0.77 and a Kappa statistic of 0.65. Notably, the harmonic model had the lowest accuracy with an overall accuracy of 0.37 and a Kappa statistic of 0.06. Figure 12 shows F1 score, precision, and recall for the wetland class and SVM models.

Model	RF OA	RF Kappa	SVM OA	SVM Kappa
Harmonic only	0.33	0.01	0.37	0.06
Medians only	0.57	0.35	0.59	0.38
Terrain only	0.85	0.78	0.77	0.65
Harmonic and Terrain	0.83	0.75	0.71	0.56
Median and Terrain	0.84	0.76	0.80	0.70
All Variables	0.83	0.74	0.77	0.65

Table 9. Overall accuracy (OA) and Kappa statistics for random forest (RF) and support vector machine (SVM) classifications.
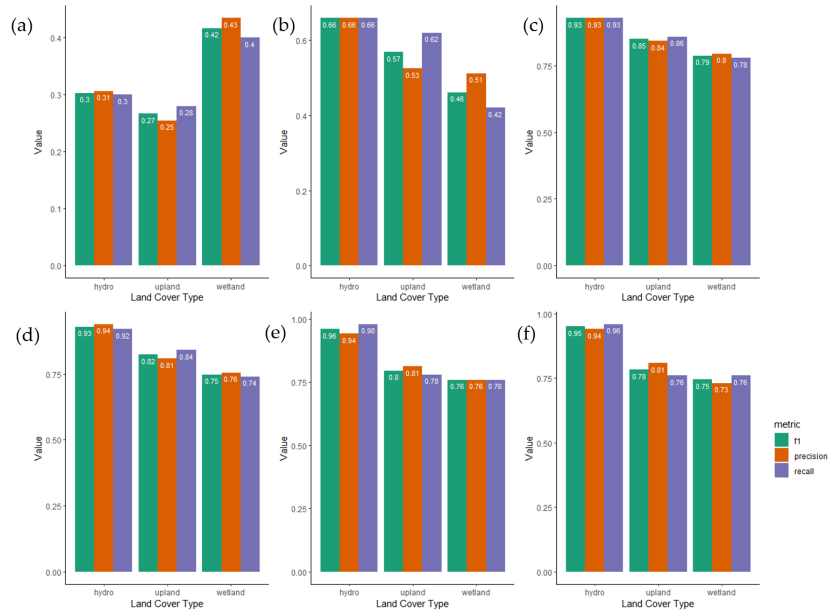
Figure 11. Random forest F1 score, precision, and recall by class when using(a) SAR harmonic variables, (b) SAR medians, (c) terrain variables, (d) harmonic and terrain variables, (e) median and terrain variables, and (f) all variables.
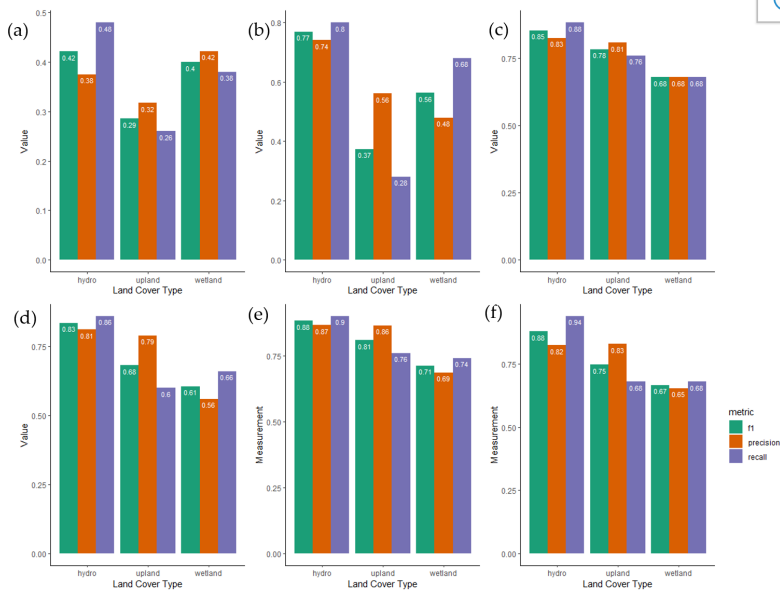
Figure 12. Support vector machine F1 score, precision, and recall by class when using (a) SAR harmonic variables, (b) SAR medians, (c) terrain variables, (d) harmonic and terrain variables, (e) median and terrain variables, and (f) all variables.

3.2 Essex County

The distribution of the VV and VH harmonic coefficients for Essex county are provided in Figure 13. Visually, outliers can be seen in all land cover types in the VV amplitude and VH amplitude datasets. While there are some visual variations between the classes in each coefficient, the VH phase angle boxplots are strikingly similar. This is statistically evident in Table 10, as the VH phase angle p-value is 0.92. VH amplitude also does not meet the threshold for statistical significance at 0.13, however, both VV phase angle and VV amplitude have statistically significant p-values. This is an improvement from Bennington county, where none of the harmonic variables reach statistical significance
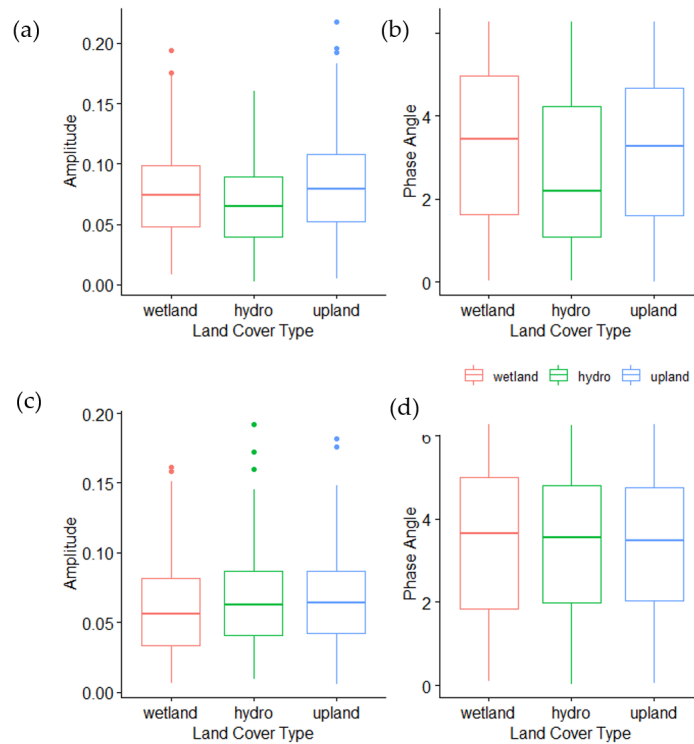
Figure 13. Boxplots of distribution for (a) VV Amplitude, (b) VV Phase, (c) VH Amplitude, and (d) VH Phase by mapped class.

Variable	*p*-value
VV Amplitude	0.00
VV Phase	0.00
VH Amplitude	0.13
VH Phase	0.92

Table 10. Resulting p-values from Kruskal-Wallis test for harmonic variables.

The boxplots for Essex county show in Figure 14 show consistent outliers for each land cover type in each model, particularly a low value in the hydrological land cover type. The Kruskal-Wallis p-values for these varied, as shown in Table 11. Profile curvature and plan curvature were not statistically significant, with p-values of 0.74 and 0.72, respectively. However, the other four variables had low enough values to reach a level of statistical significance. This is a similar case to the previous Chittenden County, where both profile curvature and plan curvature presented as statistically insignificant. This is worth noting as one of the similarities in the performance of these measures across different geographical regions.
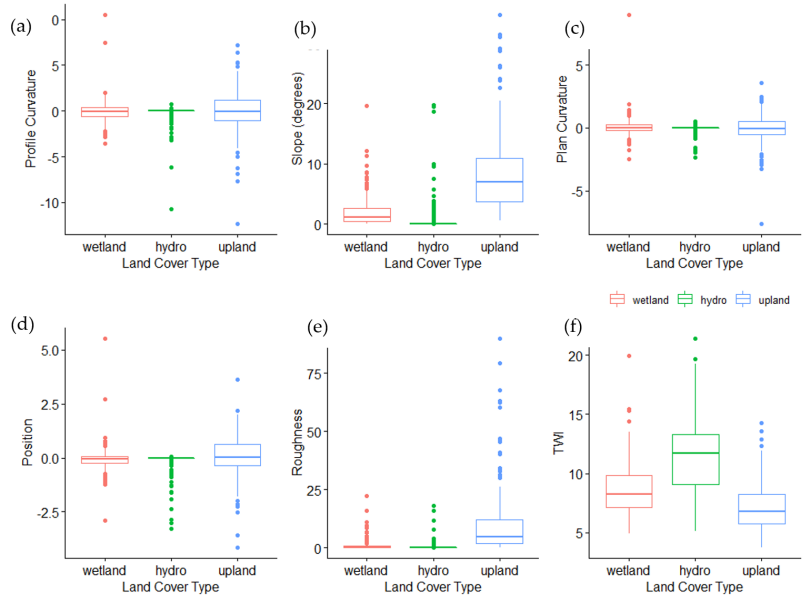
Figure 14. Boxplots of distribution for (a) Profile Curvature, (b) Slope, (c) Plan Curvature, (d) Position, (e) Roughness, and (f) TWI by mapped class.

Variable	*p*-value
Profile Curvature	0.74
Slope	0.00
Plan Curvature	0.73
Position	0.00
Roughness	0.00
TWI	0.00

Table 11. Resulting p-values from Kruskal-Wallis test for terrain variables.

The random forest and support vector machine models were run for the six model combinations in Essex County. The overall accuracies and Kappa statistics for those models are shown in Table 12. Referencing the random forest model, all three combined models had the highest accuracies, with overall accuracies of 0.81 and Kappa statistics of 0.71. Of the models

that only took one type of variable, the terrain variables outperformed the others with an overall accuracy of 0.85 and a Kappa statistic of 0.78. The harmonic model performed the worst, as in other counties, with an overall accuracy of 0.45 and a Kappa statistic of 0.17. The F1 score, precision, and recall are shown in Figure 15. Once again, the SVM models had slightly lower accuracies in most models. The model that included the median and terrain variables had the highest of the SVM accuracies, with an overall accuracy of 0.81 and Kappa statistic of 0.72. Of the models that only took one dataset, the terrain only model again had the highest with an overall accuracy of 0.74 and a Kappa statistic of 0.60. Figure 16 shows the F1 score, precision, and recall for the SVM models.
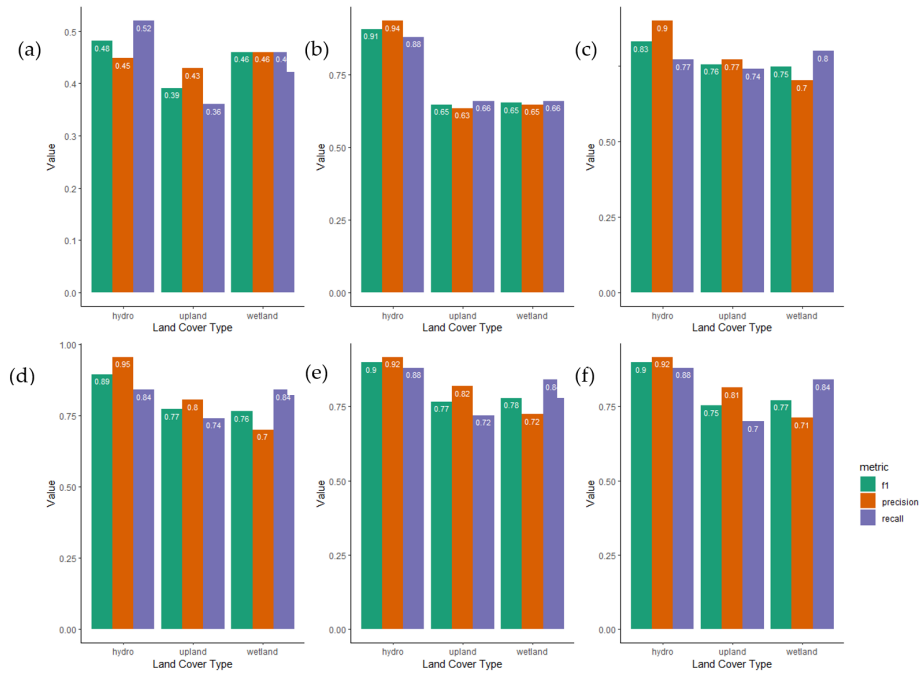

Figure 15. Random forest F1 score, precision, and recall by class when using (a) SAR harmonic variables, (b) SAR medians, (c) terrain variables, (d) harmonic and terrain variables, (e) median and terrain variables, and (f) all variables.
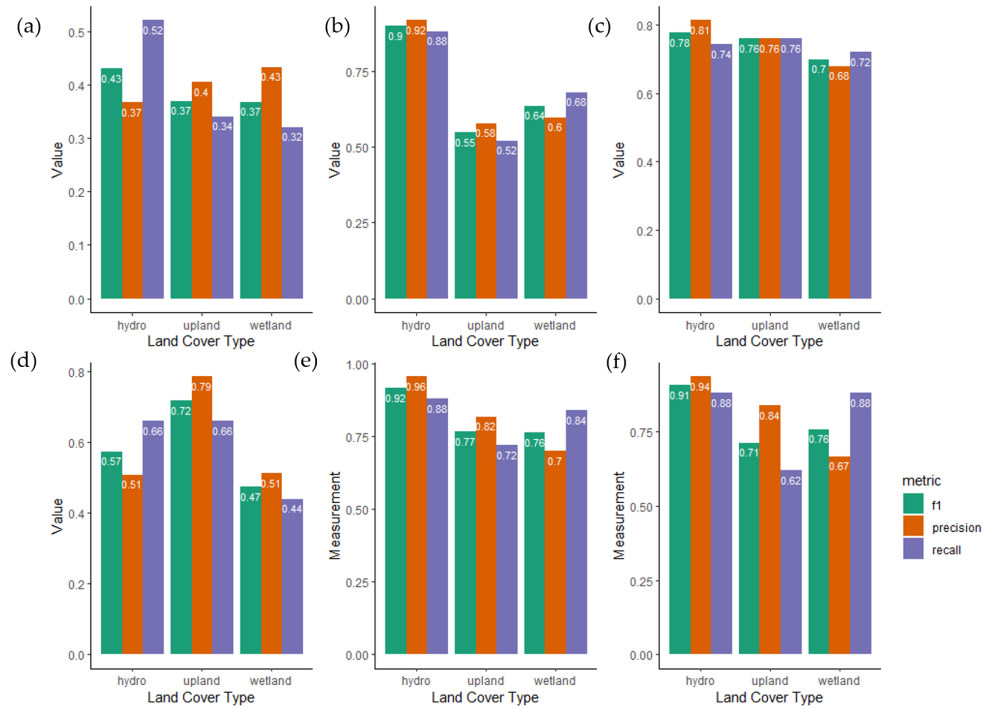
Figure 16. Support vector machine F1 score, precision, and recall by class when using (a) SAR harmonic variables, (b) SAR medians, (c) terrain variables, (d) harmonic and terrain variables, (e) median and terrain variables, and (f) all variables.

Model	RF OA	RF Kappa	SVM OA	SVM Kappa
Harmonic only	0.45	0.17	0.39	0.09
Medians only	0.73	0.60	0.69	0.54
Terrain only	0.77	0.65	0.74	0.60
Harmonic and Terrain	0.81	0.71	0.59	0.60
Median and Terrain	0.81	0.72	0.81	0.72
All Variables	0.81	0.71	0.79	0.69

Table 12. Overall accuracy (OA) and Kappa statistics for random forest (RF) and support vector machine (SVM) classifications

4. Discussion

  The exploratory data analysis provided insight into the significance of the predictor variables. Slope, slope position, roughness, and TWI had the lowest p-values among terrain variables across the board. The VV amplitude, VV phase angle, VH amplitude, and VH phase angle variables had inconsistent significance values across the board. Results suggest that the exploratory data analysis is limited in that it only considers p-values, as it is just one of many measures that can support a hypothesis [39]. Many variables had insignificant p-values but still had sufficient classification metrics, while others had statistically significant p-values but were unable to accurately classify the data. An example of the former is the terrain variables in Chittenden county, which had the highest overall accuracy and Kappa statistic for the terrain classification, yet two of the six variables did not reach statistical significance. In most cases higher accuracies were associated with terrain values—they consistently had the highest accuracies across the board when comparing only one variable. The reason for this should be investigated further in future studies but could potentially be connected to the aforementioned glacial activity in Vermont that heavily shapes its wetland landscapes.

  Interestingly the SAR data was of little value for this study. The SAR medians consistently outperformed the SAR harmonic data, indicating that a harmonic time series may not be appropriate for this dataset. This also indicates that there may be a strong difference in backscatter coefficients across the four seasons. Notably, the SAR imagery in this study was subject to the preprocessing techniques of Google Earth Engine, therefore making it harder to make adjustments to these initial values and a limitation of this study. In future research, it would be recommended to smooth the SAR data to reduce radar speckle. Other studies had noted an increase in accuracy when incorporating SAR data into optical and/or terrain datasets, however they did not assess SAR medians or SAR harmonic coefficients separately [18]. An additional consideration is the specifications of this particular satellite. For instance, a sensor with higher spatial resolution may present more accuracy when detecting smaller wetlands, or perhaps the X-band, which tends to be more sensitive than the larger C-band, could better detect the presence of water. Furthermore, this study is limited in that only VV and VH polarizations were used. This is due to the fact that these are the only polarizations collected over this area, however there is information content in HV and HH polarization that could be used in the model, as HH in particular has been noted as being effective for identifying waterbodies [17,23]. While the terrain

dataset yielded the highest classification results, it's worth noting that it did not have the seasonal dimension to it that the SAR data had. While for some measures, such as slope, there would not be much notable seasonal change, for other variables, such as TWI, these values could change greatly. This study also sought to explore the use and potential issues surrounding VSWI dataset. One potential issue with this data set is that it is modelled, rather than collected at the ground level. While this likely serves as a quicker and more cost-effective method of mapping wetlands, it also restricts this study to the accuracy of the modelled data. Lastly, when considering the VSWI dataset, it is possible that there is a mismatch in the spatial resolutions of that data and the SAR data. This is illustrated in Figure 17, as it shows how sometimes an entire wetland only spans one or two cells. This reiterates the need for higher quality, open access SAR data.
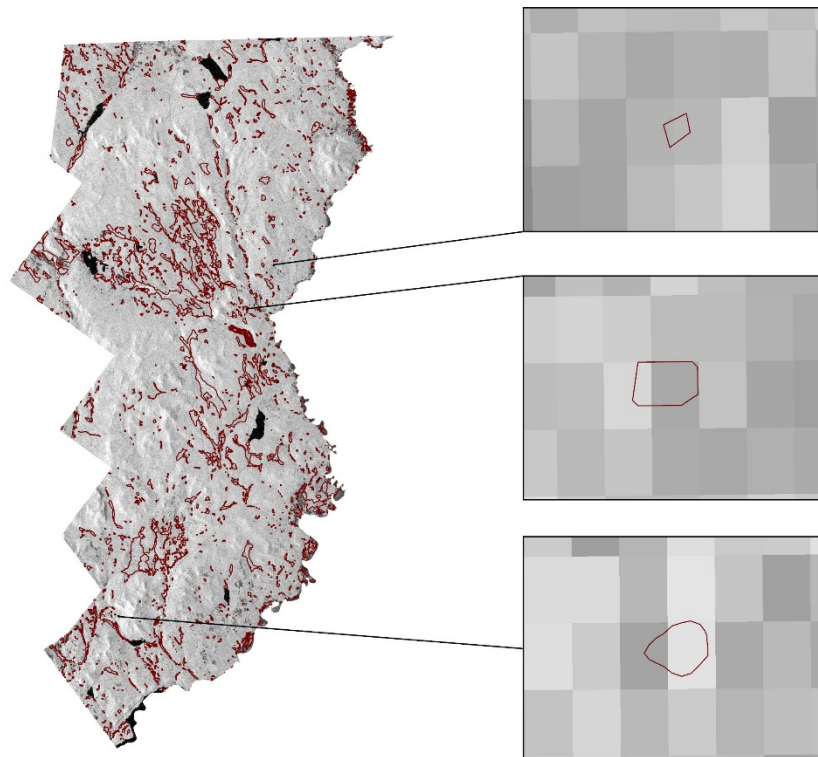

Figure 17. Essex county palustrine wetlands overlayed on a VH SAR image, highlighting three particularly small wetland areas

Additionally, the study would benefit from being performed over a larger area in order to obtain a larger sample size. There were 600 total samples per county with 200 per class—this

number was limited simply by the sheer number of points that could be collected within the palustrine wetlands. One potential option would be to have a statewide study, however that would fail to account for the unique geographic landscapes across Vermont. These are potentially worth maintaining considering the differences in p-values for the same variable across different counties. Another option would be to break the study up by its physiographic regions outlined in Figure 1 as opposed to counties. This may provide a larger area to collect more sample points, while still taking into consideration these unique landscapes. From a regulatory standpoint, this study adds to the inventory of literature that highlights the need for continuously updated, accurate, and transparent wetland data. The data and methods provided in this study are completely open source, meaning they can be conducted free of charge by any individual or organization, however the accuracy metrics from the results need to become more robust before using this methodology at a government level. Wetlands are a necessary part of our ecosystems and communities, and being able to accurately identify their locations to assess loss is paramount. Overall, this study serves as a primer for expanding the use of SAR and terrain variables for the identification of palustrine wetlands.

5. Conclusion
In considering the effects and implications of this study it is necessary to consider the original goals of this study were to (1) assess measures derived from a time series of Sentinel-1 C-band SAR for mapping palustrine wetlands and differentiating them from uplands and water bodies, (2) compare harmonic regression and seasonal aggregating techniques for summarizing a SAR time series for palustrine wetland mapping, and (3) assess DEM-derived variables for palustrine wetland mapping. The results from this study provided insight into the usefulness of SAR and DEM data for wetland mapping and the respective conclusions include the following:

1. SAR variables derived from a harmonic time series could not accurately identify wetlands alone—only when combined with terrain variables could they achieve moderate classification accuracies.
2. Seasonal aggregation in the form of seasonal medians provided better classifications accuracies than seasonal time series variables when used as the sole predictor variables.

3. DEM-derived terrain variables provided the highest accuracy of all three datasets, not including the combined dataset models. The terrain variables coupled with the SAR datasets consistently yielded higher accuracies than the SAR datasets alone.

References

1. Kingsford, R.; Basset, A.; Jackson, L. Wetlands: Conservation's Poor Cousins. Aquatic Conservation: Marine and Freshwater Ecosystems 2016, 26, 892–916.

2. Davidson, N. How Much Wetland Has the World Lost? Long-Term and Recent Trends in Global Wetland Area. Marine and Freshwater Research 2014, 65, 934–941.

3. Cantonati; Poikane; Pringle; Stevens; Turak; Heino; Richardson; Bolpagni; Borrini; Cid; et al. Characteristics, Main Impacts, and Stewardship of Natural and Artificial Freshwater Environments: Consequences for Biodiversity Conservation. Water 2020, 12, 260, doi:10.3390/w12010260.

4. Clarkson, B.R.; Ausseil, A.-G.E.; Gerbeaux, P. Wetland Ecosystem Services. 11.

5. Mitsch, W.J.; Bernal, B.; Hernandez, M.E. Ecosystem Services of Wetlands. International Journal of Biodiversity Science, Ecosystem Services & Management 2015, 11, 1–4, doi:10.1080/21513732.2015.1006250.

6. Ghermandi, A.; van den Bergh, J.C.J.M.; Brander, L.M.; de Groot, H.L.F.; Nunes, P.A.L.D. Values of Natural and Human-Made Wetlands: A Meta-Analysis: VALUES OF NATURAL AND HUMAN-MADE WETLAND. Water Resour. Res. 2010, 46, doi:10.1029/2010WR009071.

7. Hu, S.; Niu, Z.; Chen, Y.; Li, L.; Zhang, H. Global Wetlands: Potential Distribution, Wetland Loss, and Status. Science of The Total Environment 2017, 586, 319–327, doi:10.1016/j.scitotenv.2017.02.001.

8. Gallant, A. The Challenges of Remote Monitoring of Wetlands. Remote Sensing 2015, 7, 10938–10950.

9. Wright, C.; Gallant, A. Improved Wetland Remote Sensing in Yellowstone National Park Using Classification Trees to Combine TM Imagery and Ancillary Environmental Data. Remote Sensing of Environment 2006, 107, 582–605.

10. Maxwell, A.; Warner, T.; Strager, M. Predicting Palustrine Wetland Probability Using Random Forest Machine Learning and Digital Elevation Data-Derived Terrain Variables. Photogrammetric Engineering and Remote Sensing 2016, 8, 437–447.

11. Junk, W.J.; An, S.; Finlayson, C.M.; Gopal, B.; Květ, J.; Mitchell, S.A.; Mitsch, W.J.; Robarts, R.D. Current State of Knowledge Regarding the World's Wetlands and Their Future under Global Climate Change: A Synthesis. Aquat Sci 2013, 75, 151–167, doi:10.1007/s00027-012-0278-z.

12. Cowardin, L.M.; Carter, V.; Golet, F.C.; LaRoe, E.T. Classification of Wetlands and Deepwater Habitats of the United States; U.S. Fish and Wildlife Service, 1979;

13. Sharpe, P.; Kneipp, G.; Forger, A. Comparison of Alternative Approaches for Wetlands Mapping: A Case Study from Three U.S. National Parks, Wetlands 2016, 36, 547–556.

14. Tiner, R.W. NWI Maps: What They Tell Us. National Wetlands Newsletter 1997, 19, 7–12.

15. Wilkins, E.J.; Sinclair, W.; Miller, H.M.; Schuster, R.M. Does Proximity to Wetlands Matter? A Landscape-Level Analysis of the Influence of Local Wetlands on the Public's Concern for Ecosystem Services and Conservation Involvement. Wetlands 2019, 39, 1271–1280, doi:10.1007/s13157-018-1076-8.

16. Araya-Lopez, R.; Lopatin, J.; Fassnacht, F.; Hernandez, H. Monitoring Andean High Altitude Wetlands in Central Chile with Seasonal Optical Data: A Comparison Between Worldview-2 and Sentinel-2 Imagery. ISPRS Journal of Photogrammetry and Remote Sensing 2018, 145, 213–224.

17. Amani, M.; Mahdavi, S.; Afshar, M.; Brisco, B.; Huang, W.; Mirzadeh, S.M.J.; White, L.; Banks, S.; Montgomery, J.; Hopkinson, C. Canadian Wetland Inventory Using Google Earth Engine: The First Map and Preliminary Results. Remote Sensing 2019, 11, 842.

18. Corcoran, J.; Knight, J.; Brisco, B.; Kaya, S.; Cull, A.; Murnaghan, K. The Integration of Optical, Topographic, and Radar Data for Wetland Mapping in Northern Minnesota. Canadian Journal of Remote Sensing 2012, 37, 564–582.

19. Ozesmi, S.L.; Bauer, M.B. Satellite remote sensing of wetlands. Wetlands Ecology and Management 2002, 10, 381–402.

20. Hird, J.; DeLancey, E.R.; McDermid, G.J.; Kariyeva, J. Google Earth Engine, Open-Access Satellite Data, and Machine Learning in Support of Large-Area Probabilistic Wetland Mapping. Remote Sensing 2017, 9, 1315.

21. Vincent, P.; Bourbigot, M.; Johnsen, H.; Piantanida, R. Sentinel-1 Product Specification; European Space Agency, 2020;

22. Zhang, M.; Li, Z.; Bangsen, T.; Zhou, J.; Tang, P. The Backscattering Characteristics of Wetland Vegetation and Water-Level Change Detection Using Multi-Mode SAR: A Case Study. International Journal of Applied Earth Observation and Geoinformation 2015, 45, 1–13.

23. Mahdianpari, M.; Salehi, B.; Mohammadimanesh, M.; Homayouni, S.; Gill, E. The First Wetland Inventory Map of Newfoundland at a Spatial Resolution of 10 m Using Sentinel-1 and Sentinel-2 Data on the Google Earth Engine Cloud Computing Platform. Remote Sensing 2018, 11, 43.

24. Baghdadi, N.; Bernier, M.; Gauthier, R.; Neeson, I. Evaluation of C-Band SAR Data for Wetlands Mapping. International Journal of Remote Sensing 2001, 22, 71–88.

25. Agency, E.S. Sentinel-1 SAR Technical Guide; European Space Agency, 2020;

26. Schlaffer, S.; Chini, M.; Dettmering, D.; Wagner, W. Mapping Wetlands in Zambia Using Seasonal Backscatter Signatures Derived from ENVISAT ASAR Time Series. Remote Sensing 2016, 8, 402, doi:10.3390/rs8050402.

27. Maleki, S.; Baghdadi, N.; Soffianian, A.; El Hajj, M.; Rahdari, V. Analysis of Multi-Frequency and Multi-Polarization SAR Data for Wetland Mapping in Hamoun-e-Hirmand Wetland. International Journal of Remote Sensing 2019, 41, 2277–2302.

28. Chimner, R.A.; Bourgeau-Chavez, L.; Grelik, S.; Hribljan, J.A.; Planas Clarke, A.M.; Polk, M.H.; Lileskov, E.A.; Fuentealba, B. Mapping Mountain Peatlands and Wet Meadows Using Multi-Date, Multi-Sensor Remote Sensing in Cardillera Blanca. Peru. Wetlands 2019, 39, 1057–1067.

29. Banks, S.; White, L.; Behnamian, A.; Chen, Z.; Montpetit, B.; Brisco, B.; Pasher, J.; Duffe, J. Wetland Classification with Multi-Angle/Temporal SAR Using Random Forests. Remote Sensing 2019, 11, 670, doi:10.3390/rs11060670.

30. Kayastha, N.; Thomas, V.; Galbraith, J.; Banskota, A. Monitoring Wetland Change Using Inter-Annual Landsat Time-Series Data. Wetlands 2012.

31. Henderson, F.M.; Lewis, A.J. Radar Detection of Wetland Ecosystems: A Review. International Journal of Remote Sensing 2008, 29, 5809–5835, doi:10.1080/01431160801958405.

32. Wilusz, D.C.; Zaitchik, B.F.; Anderson, M.C.; Hain, C.R.; Yilmaz, M.T.; Mladenova, I.E. Monthly Flooded Area Classification Using Low Resolution SAR Imagery in the Sudd Wetland from 2007 to 2011. Remote Sensing of Environment 2017, 194, 205–218, doi:10.1016/j.rse.2017.03.005.

33. Cronin, T.M.; Manley, P.L.; Brachfield, S.; Manley, T.O.; Willard, D.A.; Guilbault, J.-P.; Rayburn, J.A.; Thunell, R.; Berke, M. Impacts of Post-Glacial Lake Drainage Events and Revised Chronology of the Champlain Sea Episode 13–9 Ka. Palaeogeography, Palaeoclimatology, Palaeoecology 2008, 262, 46–60.

34. Ridge, J.C.; Balco, G.; Bayless, R.L.; Beck, C.C.; Carter, L.B.; Dean, J.L.; Voytek, E.B.; Wei, J.H. The New North American Varve Chronology: A Precise Record of Southeastern Laurentide Ice Sheet Deglaciation and Climate, 18.2-12.5 Kyr BP, and Correlations with Greenland Ice Core Records. American Journal of Science 2012, 312, 685–722, doi:10.2475/07.2012.01.

35. Bureau, U.S.C. Resident Population and Net Change 2019.

36. Vermont Biophysical Regions 2018.

37. Natural Resources, V.A. Vermont Wetland Rules; State of Vermont, 2020;

38. Natural Resources, V.A. Vermont's Significant Wetlands Inventory 2020.

39. Nahm, F.S. What the P Values Really Tell Us. Korean J Pain 2017, 30, 241–242, doi:10.3344/kjp.2017.30.4.241.

40. Ramanathan, K.; Thenmozhi, M.; George, S.; Anandan, S.; Veeraraghavan, B.; Naumova, E.N.; Jeyaseelan, L. Assessing Seasonality Variation with Harmonic Regression: Accommodations for Sharp Peaks. International Journal of Environmental Research and Public Health 2020, 17, 1318.

41. Team, R.C.D. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing 2012.

42. Hyndman, R.; Athanasopoulos, G.; Bergmeir, C.; Caceres, G.; Chhay, L.; O'Hara-Wild, M.; Petropoulos, F.; Razbash, S.; Wang, E.; Yasmeen, F. 2020.

43. Evans, J.S.; Oakleaf, J.; Cushman, S. An ArcGIS Toolbox for Surface Gradient and Geomorphometric Modeling, Version 2.0-0.; 2015;

44. STRAHLER, A.N. QUANTITATIVE SLOPE ANALYSIS. GSA Bulletin 1956, 67, 571–596, doi:10.1130/0016-7606(1956)67[571:QSA]2.0.CO;2.

45. Berry, J.K. Use Surface Area for Realistic Calculations. Geoworld 2002, 15, 20–1.

46. Riley, S.J.; DeGloria, S.D.; Elliot, R. A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity. Intermountain Journal of Sciences 1999, 5, 1–4.

47. Bolstad, P.V.; Lillesand, T.M. Improved Classification of Forest Vegetation in Northern Wisconsin through a Rule-Based Combination of Soils, Terrain, and Landsat Thematic Mapper Data. Forest science (USA) 1992.

48. McNab, W.H. Terrain Shape Index: Quantifying Effect of Minor Landforms on Tree Height. Forest Science, Vol. 35(1): 91-104 1989.

49. Gessler, P.E.; Moore, I.D.; McKENZIE, N.J.; Ryan, P.J. Soil-Landscape Modelling and Spatial Prediction of Soil Attributes. International journal of geographical information systems 1995, 9, 421–432, doi:10.1080/02693799508902047.

50. Mountrakis, G.; Im, J.; Ogole, C. Support Vector Machines in Remote Sensing: A Review. ISPRS Journal of Photogrammetry and Remote Sensing 2011, 66, 247–259, doi:10.1016/j.isprsjprs.2010.11.001.

51. Breiman, L. Random Forests. Machine Learning 2001, 45, 5–32.

52. Dietrich, T. Overfitting and Undercomputing in Machine Learning. ACM Computer Surveys 1995, 27, 326–327.

53. Maxwell, A.; Warner, T.; Fang, F. Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. International Journal of Remote Sensing 2018, 39, 2784–2817.

54. Kuhn, M. Caret Package. Journal of Statistical Software 2008, 28.