

# Leveraging Google Earth Engine User Interface for Semiautomated Wetland Classification in the Great Lakes Basin at 10 m With Optical and Radar Geospatial Datasets

Vanessa L. Valenti<sup>✉</sup>, Erica C. Carcelen, Kathleen Lange, Nicholas J. Russo, and Bruce Chapman

**Abstract**—As one of the world’s largest freshwater ecosystems, the Great Lakes Basin houses thousands of acres of wetlands that support a variety of crucial ecological and environmental functions at the local, regional, and global scales. Monitoring these wetlands is critical to conservation and restoration efforts; however, current methods that rely on field monitoring are labor-intensive, costly, and often outdated. In this article, we present a graphical user interface constructed in Google Earth Engine called the Wetland Extent Tool (WET), which allows semiautomatic wetland classification according to a user-input area of interest and date range. WET conducts multisource, moderate resolution processing utilizing Landsat 8 Operational Land Imager, Sentinel-2 MultiSpectral Instrument, Sentinel-1 C-SAR, and Shuttle Radar Topography Mission (SRTM) datasets to classify wetlands in the entire Great Lakes Basin. We evaluated classification results of wetlands, uplands, and open water from May–September 2019, and tested whether SRTM elevation, slope, or the Dynamic Surface Water Extent produced the most accurate results in each Great Lake Basin in conjunction with optical indices and radar composites. We found that slope produced the most accurate classification in Lake Michigan, Huron, Superior, and Ontario, while elevation performed best in Lake Erie. Classification results averaged 86.2% overall accuracy, 70.0% wetland consumer’s accuracy, and 82.7% wetland producer’s accuracy across the Great Lakes Basin. WET leverages cloud-computing for multisource processing of moderate resolution remote sensing data, and employs a user interface in Google Earth Engine that wetland managers and conservationists can use to monitor wetland extent in the Great Lakes Basin in near real-time.

**Index Terms**—Graphical user interfaces (GUI), image classification, monitoring, optical image processing, synthetic aperture radar, satellite applications.

Manuscript received June 20, 2020; revised August 9, 2020; accepted August 27, 2020. Date of publication September 24, 2020; date of current version October 14, 2020. This work was supported by NASA through contract NNL16AA05C. (*Corresponding author*: Vanessa L. Valenti.)

Vanessa L. Valenti, Erica C. Carcelen, and Kathleen Lange are with NASA DEVELOP, Jet Propulsion Laboratory (Science Systems and Applications, Inc.), California Institute of Technology, Pasadena, CA 91125 USA (e-mail: vvalenti@ucla.edu; ericacarcelen@gmail.com; katielange19@gmail.com).

Nicholas J. Russo is with NASA DEVELOP, Jet Propulsion Laboratory (Science Systems and Applications, Inc.), California Institute of Technology, Pasadena, CA 91125 USA, and also with the Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095 USA (e-mail: nickrusso@ucla.edu).

Bruce Chapman is with Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: bruce.d.chapman@jpl.nasa.gov).

Digital Object Identifier 10.1109/JSTARS.2020.3023901

## I. INTRODUCTION

IT IS estimated that wetlands cover 3–5% of Earth’s land surface [1], [2]. They support a variety of crucial ecological and environmental functions. At local scales, wetlands provide wildlife habitat, flood and storm mitigation, coastal protection, and hydrologic connectivity [2], [3]. At broader scales, wetlands help regulate regional climate through carbon sequestration and provide critical habitat for continental and intercontinental migratory species [4]. Despite their vast benefits, wetland habitats are changing and disappearing due to agricultural and industrial development, water diversion, and changing precipitation patterns [5].

The subsequent need to monitor and restore wetlands has been recognized by local and federal governments in recent decades [7]. For example, the United States Fish and Wildlife Service maintain the National Wetlands Inventory (NWI), which monitors United States wetlands primarily through field data collection [6]. However, maintenance of this inventory requires extensive field work, labor, and financial investment, exacerbated by the often remote or logistically challenging to survey location of many wetlands. Comprehensive, timely wetland delineation maps are unattainable using only traditional *in situ* data collection methods. For these reasons, monitoring wetlands remotely via satellite data has become an attractive approach for long-term, comprehensive wetland delineation. Wetland distribution can be challenging to map as inundation can vary seasonally, and the term “wetland” encompasses a wide variety of ecosystem types that lack unifying features [4]. Additionally, wetlands fragmented by nature, forming patches in a matrix of surrounding habitat types. Unlike forests or open water, the boundaries of individual wetlands are difficult to monitor with satellite imagery because steep environmental gradients at wetland ecotones may not be identifiable at low-to-medium spatial resolution [4]. Wetlands’ highly dynamic nature and lack of unifying land cover features pose great challenges to train algorithms to map wetlands with levels of accuracy and consistency sufficient for monitoring. For this reason, knowledge of wetland extent and distribution through comprehensive monitoring programs is a frontier of research with important applications in ecosystem conservation and land management.

These concerns can be addressed by incorporating complementary optical and radar datasets to improve the delineation of spectrally similar wetland types [7]. Optical data are sensitive to the chemical and molecular structure of vegetation while radar data, such as synthetic aperture radar (SAR), are sensitive to the geometric and physical structure of herbaceous vegetation [7]. SAR data are especially useful to determine the flooding status of vegetation and remain unaffected by cloud cover or day/night conditions [7]. Synthesizing optical and radar datasets allows for a comprehensive classification of wetland delineation while mitigating the effects of cloud cover and night-fall. Traditionally, generating large-scale image composites and executing advanced classification algorithms required massive data storage capacity and high computational efficiency [2]. The collection, storage, processing, manipulation, and analysis of massive volumes of satellite data over large geographic areas through conventional desktop remote sensing processing programs is infeasible, otherwise termed the “Geo Big Data” problem [2]. Recently, the development of cloud-computing programs such as Google Earth Engine (GEE), the open-access availability of large volumes of Earth observation (EO) data archives and cloud-stored datasets, and advances in machine learning techniques have made comprehensive, large-scale wetland delineation maps attainable [2], [8], [9].

GEE provides access to satellite data on a planetary scale and offers extensive computing power for image processing and analysis [9]. The GEE platform is designed to help researchers easily disseminate their results to other researchers, policymakers, NGOs, and the general public [10]. GEE allows users to undertake cloud-based, planetary-scale computation without the technical expertise needed to use traditional supercomputers. This platform allows scientists and researchers to engage in applied Earth sciences rather than investing time into code development or algorithm execution. Wetland mapping in GEE has been limited mostly to local, regional, or national scales, with global land cover products focused largely on forests and surface water [11]–[15]. Much development remains to increase availability and access to reliable, large-scale wetland monitoring resources and programs. Recent efforts to utilize remote sensing data for wetland monitoring through cloud-computing programs have yielded significant studies advancing wetland mapping, including [2], [7], [8], [21], [24].

In this article, we employed GEE resources to map wetland extent in the Great Lakes Basin. The Great Lakes Basin (see Fig. 1) ecosystem is the largest body of fresh water in the world, holding 18% of the world’s freshwater supply [16]. This region covers 765 000 km<sup>2</sup>, spans 8 U.S. states and 1 Canadian province, and supplies fresh water to 35 million people. The Basin harbors over 2000 km<sup>2</sup> of wetlands that provide critical ecological services measured in economic, aesthetic, scientific, and recreational value [17], [18]. Wetland services such as water purification, maintenance of biodiversity, and flood control maintain the health of the region’s ecosystems [17]. However, the Basin is within one of the most industrialized regions of the world, encompassing nearly 25% of Canadian agricultural production and 7% of American farm production [18]. As a

result, over two-thirds of wetlands in the Great Lakes Basin have been drained for agriculture and industrial development, emphasizing the need for reliable wetland delineation maps [20].

There have been many wetland mapping initiatives in the Great Lakes Basin, such as the Great Lakes Coastal Wetland Inventory, the Ohio Wetland Inventory, and the Ontario Great Lakes Coastal Wetland Atlas, in addition to the NWI. However, these inventories are often outdated, disjointed, and do not provide comprehensive delineation of Great Lakes Basin’s wetlands. To date, a comprehensive, reliable, frequently updated wetland inventory has not been achieved in the Great Lakes Basin. For this reason, our objective was to create a 10-m resolution wetland delineation tool in GEE for the entire Great Lakes Basin using data from Landsat 8 Operational Land Imager (OLI), Sentinel-1 C-band SAR (C-SAR), Sentinel-2 MultiSpectral Instrument (MSI), and the Shuttle Radar Topography Mission (SRTM). In an effort to make wetland mapping accessible to users with varying programming experience, we developed a user-friendly interface within GEE to produce near real-time results of wetland delineation. The Wetland Extent Tool (WET) will allow decision-makers to easily access frequently updated, comprehensive, and accurate wetland extent maps to inform decisions regarding wetland conservation and restoration.

## II. MATERIALS AND METHODS

We created WET with a user-friendly graphical user interface (GUI) to allow land managers with varying levels of remote sensing and coding experience to map wetlands in the Great Lakes Basin in near real time. WET utilizes multisource data with a pixel-based random forest classifier to map the distribution of wetland, upland, and open water within user-specified dates and areas. WET visualizes the resulting map in the GEE Map Viewer for the user to instantly view, and outputs accuracy measures in the console to evaluate its performance.

WET is hosted in GEE, as it is well suited for large-scale analysis, such as the entire Great Lakes Basin; datasets that would normally require a large amount of storage space and processing time on a local computer are instead completed within the cloud. Additionally, GEE has extensive user interface options that can be utilized within the platform to create a comprehensive GUI. Employing advanced user interface options increases the usability of tools for users with little to no previous experience in GEE, thus increasing applications of geospatial analysis by decision-makers. Tools for geospatial analysis must be easily usable and accessible for frequent and repeatable analyses to inform management of dynamic ecosystems like wetlands.

### A. Data Acquisition

We accessed optical and radar Earth observations through the GEE data catalog, including imagery from the Landsat 8 OLI, Sentinel-2 MSI, Sentinel-1 C-SAR, and SRTM sensors, detailed in Table I. Landsat 8 OLI and Sentinel-2 MSI collect optical imagery every 16 days at a 30 m resolution and every 5 days at a 10 m resolution, respectively. Generally, Sentinel-1 is useful for wetland identification due to its sensitivity to wet

TABLE I  
REMOTE SENSING DATA ACQUIRED IN GOOGLE EARTH ENGINE AND USED AS INPUTS IN WET 2.0

Sensor Name	Source	Processing Level	Resolution (m)	Use
Landsat 8 OLI	NASA	Level 2 SR Tier 1	30	Surface reflectance
Sentinel-2 MSI	ESA	Level 2 SR	10	Surface reflectance
Sentinel-1 C-SAR	ESA	Level 1 GRD	10	Backscatter intensity
SRTM	NASA	Level 2 Version 3	30	Elevation; slope
Cropland Data Layer	USDA NASS	N/A	30	Agriculture mask
Canada Annual Crop Inventory	AAFC	N/A	30	Agriculture mask
Wetland Field Data	Michigan Tech; NRCan	N/A	N/A	Training and validation data
NALCMS	NRCan; CCRS, USGS	N/A	30	Urban mask

characteristics and vegetation structure in herbaceous areas, whereby the change from upland to vegetation, wetland, and open water typically yields changes in backscattering [19]. The Sentinel-1 radar dataset is also vital, even sometimes preferable for detecting structural characteristics of vegetation when cloudy or night conditions would obstruct optical imagery. We incorporated the SRTM digital elevation model (DEM) because it is one of the few topographical datasets in the GEE catalog that covered our study area with a moderate resolution of 30 m.

Additionally, we included ancillary datasets of land cover and land use in the U.S. and Canada, United States Geological Survey (USGS) shapefiles of the study area, and field data provided by our partners (see Table I). WET's base shapefile for the study area came from the USGS's Great Lakes Sub Basins shapefile, which divides the basin into 5 lake basins for reference. Field data, consisting of 681 points of surveyed coastal wetlands throughout the U.S. side of the Great Lakes Basin, were provided by the Michigan Technological Research Institute. These points, collected from 2017 to 2019, represent the center of a wetland at least  $40 \times 50$  m (0.2 ha) in size with information about wetland class, function, and health [21]–[23]. Additionally, Natural Resources Canada (NRCan) provided field data consisting of 125 polygons identifying wetland and open water in Ontario [24]. We buffered the points to create training polygons representing 0.1 ha, merged the two datasets, and reclassified to our three-class system. Our three-class system identifies upland, open water, and wetland, with upland including all land covers that are not wetland or open water. Although there are many different types of wetlands, we chose a simple three-class system due to limited representation of woody wetlands in the field data, and national differences in wetland classes and definitions between the USA and Canada. To supplement underrepresented classes, we digitized additional polygons based on interpretation of true color optical Global TruEarth 15 m imagery used as the satellite base layer in GEE, slope, and North American Land Change Monitoring System (NALCMS) 2015 land cover, which is the latest version of 30 m land cover data for the U.S. and Canada based on Landsat 7 imagery. We also utilized the Cropland Data Layer, created by the United States Department of Agriculture, and the Canadian Agriculture and Agri-Food Canada landcover dataset, to supplement addition field data from these datasets open water and wetland classes. The resulting 1615 polygons (539 upland, 406 open water, and 667 wetland), distributed

throughout the Great Lakes Basin, were used for the classifier training and data validation (see Fig. 1).

### B. Data Processing

All data processing was done in GEE—either within WET, or separately and uploaded as an asset. Both the radar and optical imagery accessed from the GEE data catalog were already processed to apply some atmospheric, radiometric, and geometric calibrations and corrections. We converted the radar data from decibels to obtain natural backscatter values for the VV and VH bands. A cloud masking function was applied to the optical datasets using each dataset's respective cloud QA bits.

Several masks were employed to prevent class confusion and minimize processing time. In radar data, scattering mechanisms dictate how features are characterized in the data. Scattering mechanisms describe how emitted waves interact with features on the ground before returning to the sensor, which include flat surface, rough surface, double-bounce, and volumetric scattering. The double bounce mechanism is of particular interest in this article because wetlands and urban land cover both have double bounce scattering returns. In a double bounce return, the emitted radar pulse first bounces off a flat surface, the ground, and then bounces off an object perpendicular to the ground, a tree in the case of wetlands and a building in the case of urban areas, before returning to the sensor. We masked out urban land cover using the NALCMS land cover dataset to address this issue, avoid class confusion, and improve accuracy. Agriculture can also present similar signatures to wetlands due to irrigation and growing practices, such as cranberry bogs. In order to differentiate wetlands from agriculture, we masked cropland data derived from the annual United States Department of Agriculture, National Agricultural Statistics Service (USDA NASS), and Agriculture and Agri-Food Canada (AAFC) crop inventory maps. WET accesses the crop inventory map from the year closest to the input date and includes all agriculture classes in the mask.

Our classification model used a variety of inputs and indices designed to identify parameters characteristic of wetlands. These inputs include the Tasseled Cap Wetness Greenness Difference Index (TCWGD), Modified Normalized Difference Water Index (MNDWI), Normalized Difference Vegetation Index (NDVI), Dynamic Surface Water Extent (DSWE), VV backscatter, VH

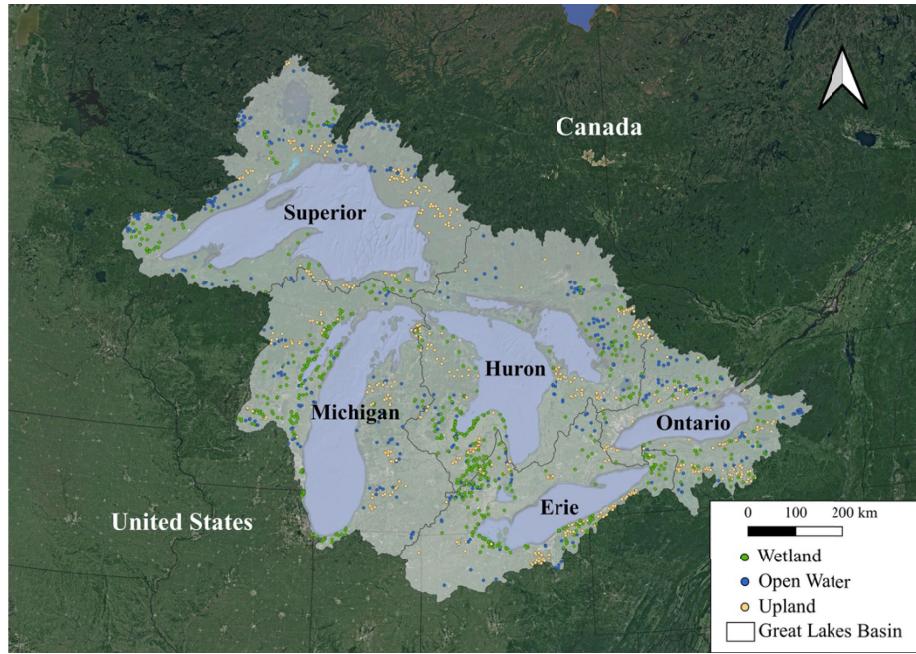


Fig. 1. Great Lakes Basin and field data distribution.

backscatter, and VV/VH ratio. Due to its finer spatial resolution and more frequent temporal resolution, our tool used Sentinel-2 MSI optical data rather than Landsat 8 OLI to calculate TCWGD, MNDWI, and NDVI. These indices are commonly utilized for wetland classification due to their ability to distinguish vegetation and water characteristics from other land cover types [25]–[27].

The optical data from Landsat 8 OLI were combined with topographic data from SRTM to calculate DSWE. Although topographic inputs have been shown to improve wetland classifications because of their relationship with water bodies, topography itself does not change frequently [28]. Meanwhile, wetland conditions and inundation change frequently, which can be captured in an index like DSWE that combines stationary topographic information with dynamic vegetation, soil, and wetness parameters [29]. DSWE, which was developed using Landsat 8 imagery, employs an algorithm that applies five decision rule-based diagnostic tests. This algorithm determines if a pixel is fully covered by water or detects inundation in the presence of nonwater land covers. The pixel is then assigned to one of six classes: not water, water (high confidence), water (moderate confidence), potential wetland, water or wetland (low confidence), or NA [9], [29]. As our three topography inputs, we incorporated DSWE as well as two traditional topographic measures of DEM elevation values and derived slope values. Radar data helped identify inundated wetlands using the VV backscatter values, VH backscatter values, and VV/VH ratio. Open water should have low backscatter values due to smooth surface scattering, and inundated wetlands should have high backscatter values due to double bounce scattering [9]. The VV/VH ratio is particularly useful for wetland classification because it is sensitive to soil moisture and corrects for terrain effects.

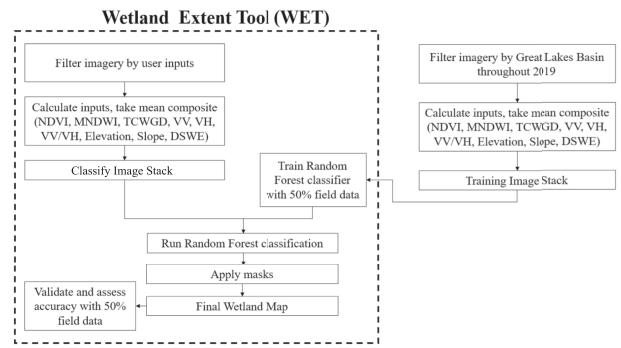


Fig. 2. Flowchart of scheme for semiautomated wetland classification.

### C. User Interface

To maximize usability, we designed a GUI with several user selections available. For WET to create a wetland map, the user must select an area of interest (AOI) and a date range (see Fig. 2). The tool then filters and crops the imagery for classification based on these parameters. Users can select from either preloaded AOIs that range from entire lake basins to specific locations like the St. Clair Delta in Michigan, or outline their own AOI using the built in GEE geometries drawing tools in the Map viewer. The input date range allows users to create maps representing any span of time, whether it be seasonal, or annual; while users can view optical indices (NDVI, MNDWI, TCWGD), and C-SAR composites on a near-daily basis (due to the high temporal resolution of Sentinel-1 and Sentinel-2), wetland classifications require at least 2–3 months of sufficient data input and cloud-free images. To better understand the output classification, the user can choose to map imagery for the

classification variables and create line graphs of variable values over time. To access results outside of GEE, the user can also use the GUI to export maps as geotiffs to their personal Google Drive. All of these options allow users to create customized wetland classification maps with ease and provide several methods to further analyze resulting maps. Additional options, such as customizing classifier settings, uploading a user's own field data points, obtaining accuracy measures detailed in Section E, and evaluating the relative importance of input variables, can be accessed within the WET code for users with more GEE JavaScript experience.

#### D. Random Forest Classification

WET conducts wetland classification using the random forest (RF) algorithm, a nonparametric machine learning technique that creates multiple decision trees based on the training data and bands of an input image stack [2], [21]. This method is frequently used in other studies for wetland mapping because it can handle datasets that stem from multiple sources, has few observations, and is not normally distributed while remaining insensitive to overfitting and noise [21], [30]. We employed the RF method on a pixel-by-pixel basis because of processing limits in GEE and it can produce satisfactory results that are not significantly different from the object-based approaches in wetland landscapes [31].

To ensure WET could classify anywhere in the Great Lakes Basin, regardless of whether there were field data present in the user-specified AOI, we incorporated two methods into the classification process (see Fig. 2). Rather than clip the field data to the user-specified AOI, we filtered the field data based on the lake basin corresponding to the user-specified area. To do this, WET calculated the centroid of the user-specified area, identified the lake basin where the centroid was located, and selected all the field data within that basin. The second method we employed was to create two separate image stacks, one for training and one for classification, that both consisted of mean composites of the inputs described in Section B. Backlogs in the GEE data catalog limited the availability of corrected Level 2 Sentinel-2 MSI data. Thus, we created the training image stack with imagery for the entire year of 2019 throughout the Great Lakes basin. The resulting training stack was imported as an asset in WET and the classification image stack was created within WET for the user-specified AOI and dates.

Field data were first randomized in GEE, and then split for training and validation samples. We employed 60% of the filtered field polygons and the values for each parameter in the training image stack to train the RF classifier. The trained RF classifier was then applied to the classification stack to create an initial classification identifying upland, open water, and wetland. We used an RF classifier with 100 trees and the default GEE settings. RF uses the decision trees it generates to determine classification by selecting the class with the most votes among all the trees. We then applied the urban and agriculture masks, automatically reclassifying these areas as upland, to produce the final wetland classification maps.

#### E. Accuracy Analysis

To evaluate the accuracy of resulting wetland maps, we used the remaining 40% of the field data polygons for validation against WET's outputs. The output wetland maps were randomly sampled to generate "predicted" values (classified as wetland) for comparison against the field "actual" values. We then generated a confusion matrix and conducted an accuracy assessment in GEE. We also computed the Kappa coefficient, consumer's accuracy, and producer's accuracy in GEE for additional statistical analysis. The overall accuracy assessment indicates what percentage of field data points were correctly classified by WET, while the Kappa coefficient evaluates the classification's performance compared to a random classification of pixels. Consumer's accuracy measures how often the class on the map will actually be present on the ground (reliability), by calculating the percentage of validation data currently classified by the tool, while producer's accuracy measures how often real features on the ground are correctly shown on the classified map, by calculating what percentage of results in the tool were confirmed by the validation data.

### III. RESULTS AND DISCUSSION

To evaluate our tool's performance and usability, we conducted several tests throughout the study area, analyzing results for each lake basin with each topographic input: DSWE, elevation, and slope to determine which variable produced the most accurate classification of wetlands. We present a statistical summary and quantitative analysis of each lake basin's classification results produced in WET, running the tool over a date range of May 25th–September 25th, 2019. We exhibit how topographic information influences results in the Lake Erie Basin, discuss misclassifications in Lake Michigan with confusion matrices, and finally consider the potential for WET to inform wetland management in the Great Lakes Basin.

#### A. Classification Results in the Great Lakes Basin

The classification results for the entire Great Lakes Basin are presented in Fig. 3, using each basin's most accurate results. The statistical summary of results for each basin's most accurate results is presented in Table II. Overall, the slope and elevation variables performed better than DSWE across the entire Great Lakes Basin. Wetland classification of Lake Ontario, Huron, Superior, and Michigan Basins was most accurate with slope as the topographic input, while Erie identified wetlands most accurately with elevation.

Table II summarizes statistics for each Lake Basin results, as well as averages across the entire Great Lakes Basin. Overall, when compared to the reserved validation dataset, our methodology achieved an average OA of 86.2% classifying wetlands, upland, and open water in the Great Lakes Basin. Consumer's accuracy averaged 70.0% of the wetlands class, 99.5% of open water, and 83.4% of uplands. Producer's accuracy averaged 82.7% of wetlands, 99.2% of open water, and 68.0% of uplands. The Kappa statistics averaged 77.7%, suggesting that our classification results were overall in high agreement with the

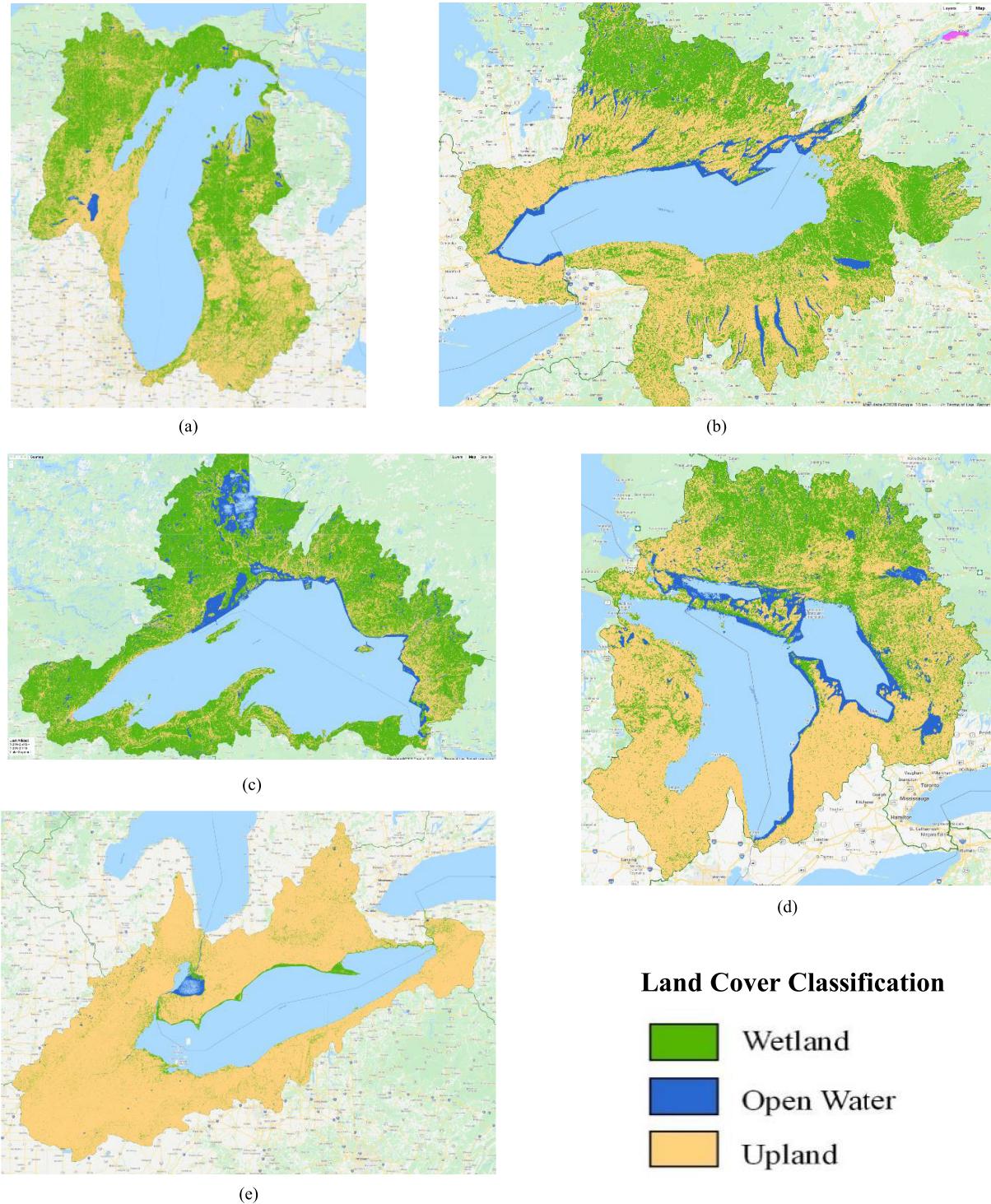


Fig. 3. Classification output of WET for each Great Lakes subbasin. (a) to (e) are Lakes Michigan, Ontario, Superior, Huron, and Erie, respectively.

validation data, even accounting for random agreement between datasets.

Lake Erie was one of the most accurately classified Basins, with a 92.1% OA, as well as the highest consumer's and producer's accuracy for the wetland class: 87.9% of all wetlands in the validation field data were correctly classified by the tool, and 97.2% of all wetlands classified by the tool were correct in the

validation data. Lake Michigan Basin classification produced an OA of 86.9%, with wetland class consumer's accuracy at 71.6% of validation wetlands correct in the results, and a producer's accuracy of 84.5%. The Lake Huron Basin classification produced an OA of 81.6% between upland, open water, and wetland classification. Huron's consumer's accuracy for wetlands was 54.8%, suggesting over half of wetlands in the validation data

TABLE II  
STATISTICAL SUMMARY OF MOST ACCURATE RESULTS FOR EACH GREAT LAKES BASIN

Lake Basin	Topographic Input	Overall Accuracy	Kappa	Consumer's Accuracy			Producer's Accuracy		
				Upland	Open Water	Wetland	Upland	Open Water	Wetland
Superior	Slope	96%	91.1%	91.4%	99.9%	84.4%	71.7%	99.9%	95.5%
Erie	Elevation	92.1%	87.2%	95.5%	99.3%	87.9%	79.9%	100%	97.2%
Huron	Slope	81.6%	71.2%	86.7%	99.8%	54.8%	75.2%	99.9%	72.2%
Michigan	Slope	86.9%	77.3%	67.9%	99.4%	71.6%	48.3%	99.9%	84.5%
Ontario	Slope	74.6%	61.6%	75.4%	99.1%	50.8%	65.3%	96.1%	64.1%
Average	N/A	86.2%	77.7%	83.4%	99.5%	70.0%	68.0%	99.2%	82.7%

TABLE III  
CONSUMER'S ACCURACY CONFUSION MATRIX FOR LAKE MICHIGAN RESULTS

Lake Michigan		Validation		
		Upland	Open Water	Wetland
Classification	Upland	67.7%	0.32%	28.3%
	Open Water	0.00%	99.4%	0.08%
	Wetland	32.3%	0.28%	71.6%

were correctly identified by the tool, and producer's accuracy for wetlands reached 72.7%, indicating most of the wetlands classified by the tool were correct in the field data. Although Lake Ontario Basin yielded the lowest OA of our results at 74.6%, wetland consumer's accuracy (50.8%), and wetland producer's accuracy (64.1%) indicates that over half of wetlands were accurately identified. In the Lake Superior Basin, results showed abundant wetland classification throughout the entire binational basin. Although statistical results indicate Superior to have produced a 96% OA, with a wetland consumer's accuracy of 84.4% and a producer's accuracy of 95.5%, we believe the Superior basin to have overclassified wetlands on the Canadian side of the Basin, leading to misleadingly high accuracies. This may be due to confusion with forested land and wetlands in the region. This confusion may be avoided in future wetland classification endeavors with the development of the Global Ecosystem Dynamics Investigation (GEDI) satellite LiDAR mission and NASA-ISRO L-band SAR (NISAR). NISAR and GEDI promise to usher in a new age of wetland mapping by revealing structural and topographical features relevant to wetland formation and

differences among wetland types that are not distinguishable with C-SAR [14], [32], [33].

#### B. Misclassification

To closely analyze and better understand the class confusion, Tables III and IV show consumer's and producer's accuracy confusion matrices for Lake Michigan. These tables break down results for each class between the classification and validation data, illustrating that the wetland class was most often misclassified as upland. Concerning consumer's accuracy, **28.3% of validation wetlands were misclassified as upland** in the results, while **32.3% of validation uplands were misclassified as wetlands** in the results. The producer's accuracy **confusion matrix** reveals that uplands were most commonly misclassified as wetlands, with just 48.2% of upland classifications correct, and most misclassification of upland was due to incorrect classification as wetland (50.8%). From these confusion matrices, we can see that **class confusion was primarily caused by upland–wetland misclassification, rather than upland–open water or wetland–open water misclassification.**

TABLE IV  
PRODUCER'S ACCURACY CONFUSION MATRIX FOR LAKE MICHIGAN RESULTS

Lake Michigan		Validation		
		Upland	Open Water	Wetland
Classification	Upland	48.2%	1.04%	50.8%
	Open Water	0.00%	99.9%	0.05%
	Wetland	15.1%	0.60%	84.3%

TABLE V  
LAKE ERIE TOPOGRAPHIC INPUT RESULTS

Variable	OA	Kappa	Consumer's Accuracy			Producer's Accuracy		
			Upland	Open	Wetland	Upland	Open	Wetland
DSWE	84.3%	75.2%	93.8%	98.8%	74.4%	65.2%	99.2%	95.4%
Slope	86.6%	78.7%	94.3%	98.5%	78.5%	69.3%	99.0%	95.9%
Elevation	92.1%	87.2%	95.5%	99.3%	87.9%	79.9%	100%	97.2%

### C. Evaluation of Topographic Input

To demonstrate the results produced when utilizing each topographic variable, Table V reports statistics for DSWE, elevation, and slope results for the Lake Erie Basin. Elevation produced the most accurate classification results, as exemplified by the higher OA, Kappa, consumer's, and producer's accuracy (see Table V). The elevation variable yielded 87.9% consumer's accuracy for the wetland class, followed by 78.5% with the slope band, and 74.4% with DSWE. Similarly, producer's accuracy decreased from 97.2% with elevation, to 95.9% with slope, and 95.4% with DSWE. This demonstrates how SRTM information produced more accurate results when identifying wetlands in the Great Lakes Basin, suggesting that raw, less refined topographic information is more useful when mapping wetlands in conjunction with optical and radar data. As a refined, highly tested, and calculated product, DSWE should not be utilized in conjunction with other highly informative geospatial data such as optical and radar data for wetland identification. Although DSWE includes topographic information in its calculation, it does not provide sufficient topographic context for wetland identification in a multisource classification.

### D. Cloud Computing Limitations, Benefits, Access, and Application

Although GEE is designed for the global analysis, user memory limitations and computation time-outs pose a significant problem to highly sophisticated methodologies or larger

scale studies. We developed several novel methods to work around GEE user memory limits and time-outs to enable WET's multisource, basin-wide classification. Through our interface, users can easily filter imagery in the GEE cloud-catalog by the desired date and area, and the tool conducts calculations and classification only according to the user's specifications to reduce the computation and user memory needed. This breaks the methodology down into small enough scales that WET is capable of conducting significant aggregations, derivations, and a thorough classification in the Great Lakes Basin region without GEE user memory limits and timeouts. Without these innovations or the GUI that executes only to the scale the user specifies, our comprehensive methodology of multisource 10 m classification is not currently possible within GEE's console, as computations and map layers time-out after 5 min [34]. In addition, other tools such as image segmentation or an object-based image analysis before RF classification were not possible within GEE limitations. We also developed novel handling of field data for training and evaluation of the RF classifier. WET filters field data to the corresponding Lake Basin in which the user has selected and trains the RF classifier on a preprocessed image stack for the entire Great Lakes Basin to allow computation within GEE's console and Map viewer user memory limit. Traditional training of the classifier that would include entire training sample overloads GEE computation limits. Despite these adaptations, errors and time-outs will occur if classifying a region larger than a single Basin, or one-year plus time periods within WET.

In addition to enabling multisource, 10 m classification within GEE's user memory and computational limits, WET has the ability to inform decision-making in wetland management and conservation. Wetland managers, conservationists, and researchers in the Great Lakes Basin who are unfamiliar with the processing and use of remote sensing data can easily access, process, and display remote sensing data through the user interface. WET then automatically applies remote sensing data to wetland dynamics at the user's local or regional needs. The separate training and classifying stack methodology also allows basin-wide classification, even if no field data are present in the area of interest—a common barrier to classifications and mapping, as training data for classifications is restricted by the amount, distribution, and quality of field methods that are timely, expensive, and not expansive [35]. Although the tool is capable of classifying outside the GLB, it is not recommended by the authors at this time as training and validation of the tool is based on data from this region and its distinct environment and ecosystems. Running the tool outside the Basin may compromise wetland identification accuracy. This development enables decision makers to use near real-time optical and radar data in monitoring and management of wetlands at the local and regional scale, where field data may not be recorded. As an interactive tool designed for those with less knowledge of geospatial data, it is imperative that users are able to complete processing and visualize results easily and quickly within the tool. Similar studies and GEE projects have produced results at similar scales, study areas, inputs, and overall method as this study, however all have relied on length computations that require exporting to the user's cloud-storage, then subsequent visualization and analysis outside of/reimported into GEE. Our tool successfully classifies wetlands throughout the GLB at 10 m within the GEE console's and Map user memory and commutation limits, allowing users unfamiliar with remote sensing and GIS data to visualize and analyze their results within the tool's interface.

Our efforts to advance application and accessibility of geospatial datasets and cloud-computing GEE services should be considered within the limitations of GEE outlined and the results and issues analyzed. Traditional methods of training and classifying remote sensing data at this scale and extent cannot be complete within the GEE console without adaptations. In the future, options to increase user memory could advance efforts to conduct sophisticated multisource, higher resolution analyses in GEE. Efforts to develop GUIs in GEE should continue to be explored, as combining GEE's expansive catalog of analysis-ready, cloud-stored data with its interface options increases access, use, and application of remote sensing data. WET provides near real-time monitoring by leveraging the GEE cloud-stored, preprocessed, analysis-ready datasets that are frequently updated and automatically fed into the tool for processing and classification. As a result, the tool is self-sustaining and will consistently provide timely wetland delineation maps in the Great Lakes Basin.

#### IV. CONCLUSION

We utilized GEE to successfully develop WET, a tool that delineates wetlands according to a user-input area and date

range by filtering, processing, and classifying geospatial data for research and management of wetlands in the entire Great Lakes Basin. We constructed a process whereby the tool classifies user-input areas trained on a separate dataset in order to delineate wetlands anywhere in the Basin regardless of field data presence over local and regional applications. WET also effectively balances GEE cloud-computing and user memory limits with multisource 10 m classification by tailoring analysis to the user's specifications. Our results show that the tool delineates wetlands, upland, and open water to a highly accurate degree in most of the Great Lakes Basin, and was relatively successful at identifying wetlands when compared to reserved validation field data, averaging 86.2% OA, 70.0% wetland consumer's accuracy, and 82.7% wetland producer's accuracy. SRTM information performed better than DSWE in the classification across the entire Great Lakes Basin, with slope generating the most accurate results in Michigan, Ontario, Superior, and Huron, while elevation succeeded best in Erie. Although DSWE is a highly informative refined product, it is not well suited as a classification input in multisource classification as applied in this article. The tool failed to identify wetlands in the Lake Superior Basin, likely due to insufficient wetland training data and a forest dominated landscape. By leveraging GEE cloud-computing and user interface options to construct WET, we increased the usability of remote sensing data for applications of wetland monitoring by those less familiar with satellite data processing and analysis. Decision-makers can utilize WET to inform wetland monitoring and management, while remote sensing and wetland researchers can easily repeat and revise WET's inputs and methodology in the GEE JavaScript API for further exploration.

#### ACKNOWLEDGMENT

The authors would like to thank the reviewers for their time, suggestions, and improvements to this article. They thank NASA DEVELOP National Program for supporting this project. They also thank L. Bourgeau-Chavez and B. Brisco for providing field data used in WET and C. Byles and B. Holt for their insight throughout the project and offering feedback on this article. They thank NASA DEVELOP team members at the Jet Propulsion Laboratory for their past contributions to a similar project: E. O'Connor, M. Ferriter, A. Lin, and C. Notto.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration. The Wetland Extent Tool code will become publicly available within 3-9 months on the NASA DEVELOP github: <https://github.com/NASA-DEVELOP>. NASA DEVELOP project page containing additional information: <https://develop.larc.nasa.gov/2020/spring/GreatLakesWaterII.html>

#### REFERENCES

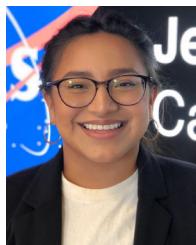
- [1] R. W. Tiner, M. W. Lang, V. V. Klemas, and C. A. Johnston, *Remote Sensing of Wetlands: Applications and Advances*. Boca Raton, FL, USA: CRC Press, 2015 [Online]. Available: [https://openprairie.sdstate.edu/nrm\\_book/2/](https://openprairie.sdstate.edu/nrm_book/2/)

- [2] M. Mahdianpari, B. Salehi, F. Mohammadimaneesh, S. Homayouni, and E. Gill, "The first wetland inventory map of Newfoundland at a spatial resolution of 10m using Sentinel-1 and Sentinel-2 data on the Google Earth engine cloud computing platform," *Remote Sens.*, vol. 11, no. 1, Dec. 2018, Art. no. 43.
- [3] W. J. Mitsch, B. Bernal, A. M. Nahlik, U. Mander, L. Zhang, and C. J. Anderson, "Wetlands, carbon, and climate change," *Landscape Ecol.*, vol. 28, pp. 583–597, Jun. 2012.
- [4] A. Gallant, "The challenges of remote monitoring wetlands," *Remote Sens.*, vol. 7, no. 8, pp. 10938–10950, Aug. 2015.
- [5] J. M. Corcoran, J. F. Knight, and A. L. Gallant, "Influence of multi-source and multi-temporal remotely sensed and ancillary data on the accuracy of random forest classification of wetlands in northern Minnesota," *Remote Sens.*, vol. 5, no. 7, pp. 3212–3238, Jul. 2013.
- [6] "NWI Program Overview," US Fish & Wildlife Service, May 2020. [Online]. Available: <https://www.fws.gov/wetlands/nwi/Overview.html>
- [7] M. Mahdianpari *et al.*, "Big data for a big country: The first generation of Canadian wetland inventory map at a spatial resolution of 10-m using Sentinel-1 and Sentinel-2 data on the Google Earth engine cloud computing platform," *Can. J. Remote Sens.*, vol. 46, no. 1, pp. 15–33, Jan. 2020.
- [8] J. Hird *et al.*, "Google Earthengine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping," *Remote Sens.*, vol. 9, no. 12, Dec. 2017, Art. no. 1315.
- [9] S. K. Ahmad, F. Hossain, H. Eldardiry, and T. M. Pavelsky, "A fusion approach for water area classification using visible, near infrared and synthetic aperture radar for south Asian conditions," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2471–2480, Apr. 2020.
- [10] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, Dec. 2017.
- [11] M. C. Hansen *et al.*, "High-resolution global maps of 21st-century forest cover change," *Science*, vol. 342, no. 6160, pp. 850–853, 2013.
- [12] G. Donchyts *et al.*, "Earth's surface water change over the past 30 years," *Nature Climate Change*, vol. 6, no. 9, pp. 810–813, 2016.
- [13] J.-F. Pekel *et al.*, "High-resolution mapping of global surface water and its long-term changes," *Nature*, vol. 540, no. 7633, pp. 418–422, 2016.
- [14] Q. Wu *et al.*, "Integrating LiDAR data and multi-temporal aerial imagery to map wetland inundation dynamics using Google Earth engine," *Remote Sens. Environ.*, vol. 228, pp. 1–13, 2019.
- [15] X. Wang *et al.*, "Tracking annual changes of coastal tidal flats in China during 1986–2016 through analyses of Landsat images with Google Earth engine," *Remote Sens. Environ.*, vol. 238, 2020, Art. no. 110987.
- [16] "Great Lakes Basin Ecosystem," US Fish & Wildlife Service, 2015. [Online]. Available: <https://www.fws.gov/midwest/fisheries/library/fact-GreatLksBasinEco.pdf>
- [17] G. Krantzberg and C. De Boer, "A valuation of ecological services in the Laurentian Great Lakes Basin with an emphasis on Canada," *J. Amer. Water Works Assoc.*, vol. 100, no. 6, pp. 100–111, Jun. 2008.
- [18] N. P. Danz *et al.*, "Integrated measures of anthropogenic stress in the U.S. Great Lakes Basin," *Environ. Manage.*, vol. 39, pp. 631–647, Mar. 2007.
- [19] D. Mohammed and B. Brisco, "Wetland monitoring and mapping using synthetic aperture radar," *Wetlands Manage. Assessing Risk Sustain. Solutions*, 2019, doi:[10.5772/intechopen.80224](https://doi.org/10.5772/intechopen.80224).
- [20] "Where are great lakes coastal wetlands?," United States Environmental Protection Agency, Washington, DC, USA, Feb. 2019. [Online]. Available: <https://www.epa.gov/great-lakes-monitoring/where-are-great-lakes-coastal-wetlands>
- [21] L. Bourgeau-Chavez *et al.*, "Development of a bi-national great lakes coastal wetland and land use map using three-season PALSAR and Landsat imagery," *Remote Sens.*, vol. 7, no. 7, pp. 8655–8682, Jul. 2015.
- [22] L. Bourgeau-Chavez *et al.*, "Mapping invasive *Phragmites australis* in the coastal Great lakes with ALOS PALSAR satellite imagery for decision support," *J. Great Lakes Res.*, vol. 39, pp. 65–77, Dec. 2012.
- [23] L. L. Bourgeau-Chavez, K. P. Kowalski, M. J. Battaglia, and A. F. Poley, "Land cover map including wetlands and invasive *Phragmites* circa 2017," U.S. Geological Survey Data Release, Nov. 2019, doi:[10.5066/P9OX21T6](https://doi.org/10.5066/P9OX21T6).
- [24] M. Mahdianpari *et al.*, "Large area big data analytics: the Canadian wetland inventory map at 10-m resolution using Senintel-1 and Sentinel-2 data and improved ground data on the Google Earth engine cloud computing platform," submitted for publication.
- [25] T. Shi and H. Xu, "Derivation of tasseled cap transformation coefficients for Sentinel-2 MSI at-sensor reflectance data," *IEEE J. Sel. Topics Appl.*, vol. 12, no. 10, pp. 4038–4048, Oct. 2019.
- [26] B. Gao, "NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.
- [27] J. W. Rouse Jr., R. H. Haas, D. W. Deering, J. A. Schell, and J. C. Harlan, "Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation," Texas A&M University, College Station, TX, USA, Final Rep. RSC 1978-2, Oct. 1973. [Online]. Available: <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740004927.pdf>
- [28] J. B. Bwangoy, M. C. Hansen, D. P. Roy, G. De Grandi, and C. O. Justice, "Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices," *Remote Sens. Environ.*, vol. 114, pp. 73–86, Aug. 2009.
- [29] J. Jones, "Improved automated detection of subpixel-scale inundation—Revised dynamic surface water extent (DSWE) partial surface water tests," *Remote Sens.*, vol. 11, no. 4, Feb. 2019, Art. no. 374.
- [30] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sánchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm.*, vol. 67, pp. 95–104, Dec. 2011.
- [31] T. M. Berhane *et al.*, "Comparing pixel- and object-based approaches in effectively classifying wetland-dominated landscapes," *Remote Sens.*, vol. 10, no. 1, 2018, Art. no. 46.
- [32] R. Dubayah *et al.*, "The global ecosystem dynamics investigation: High-resolution laser ranging of the earth's forests and topography," *Sci. Remote Sens.*, vol. 1, 2020, Art. no. 100002.
- [33] R. Valbuena *et al.*, "Standardizing ecosystem morphological traits from 3D information sources," *Trends Ecol. Evol.*, vol. 35, no. 8, pp. 656–667, 2020.
- [34] "Debugging guide | Google Earth engine | Google developers," Google, 2020. [Online]. Available: [developers.google.com/earth-engine/debugging](https://developers.google.com/earth-engine/debugging)
- [35] A. Juel, G. B. Groom, J.-C. Svenning, and R. Erjnaes, "Spatial application of Random Forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 42, pp. 106–114, Jul. 2015.
- [36] J. D. Midwood and P. Chow-Fraser, "Connecting coastal marshes using movements of resident and migratory fishes," *Wetlands*, vol. 35, pp. 69–79, 2015.
- [37] T. M. Berhane *et al.*, "Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory," *Remote Sens.*, vol. 10, no. 4, 2018, Art. no. 580.
- [38] T. G. Howard, A. L. Feldmann, E. Spencer, R. R. Ring, K. A. Perkins, and J. Corser, "Wetland monitoring for Lake Ontario adaptive management," New York Natural Heritage Program, Albany, NY, USA, 2016.



**Vanessa L. Valenti** received the B.A. degree in geography/environmental studies from the University of California, Los Angeles, CA, USA, in 2019.

She has researched ocean water quality in the Mesoamerican Barrier Reef in the past with NASA DEVELOP (Hampton, VA, USA) at the Jet Propulsion Laboratory, Pasadena, CA, USA. She previously developed the Optical Reef and Coastal Area Assessment tool in GEE that monitors turbidity, chlorophyll-a, and colored dissolved organic matter. Her research interests include applications of remote sensing and geospatial data, cloud-computing, and physical geography.



**Erica C. Carcelen** received the B.S. degree in biology and environmental science from Rhodes College, Memphis, TN, USA, in 2017 and the M.Sc. degree in geographic information science for development and environment from Clark University, Worcester, MA, USA, in 2019.

In the past, she has participated in animal behavior research with the Memphis Zoo with a Rhodes College summer fellowship, and field research with the La Selva Biological Station with the National Science Foundation Research Experience for Undergraduates.

In 2019, she researched the feasibility of relating MODIS evapotranspiration to land cover for water resource management in Panama with NASA DEVELOP in Pasadena at JPL, Hampton, VA, USA. Her research interests include environmental applications of GIS and remote sensing, natural resource management, and conservation.



**Kathleen Lange** received the B.S. degree in environmental earth science from the California Polytechnic State University, San Luis Obispo, CA, USA, in 2018.

She is currently researching permafrost subsidence and thermokarst formation effects on transportation and infrastructure in Alaska with NASA DEVELOP in Pasadena at JPL, Hampton, VA, USA. Her current research interests include water quality and resource management, applications of remote sensing, and geomorphology.



**Nicholas J. Russo** received the B.S. degree in ecology and evolutionary biology from the University of Connecticut, Storrs, CT, USA, in 2018. He is currently working toward the Ph.D. degree in ecology and evolutionary biology with the University of California, Los Angeles, CA, USA.

His current research interests include movement ecology, ornithology, seed dispersal, and applications of remote sensing, especially in the African tropics.



**Bruce Chapman** received the A.B. degree in both physics and astronomy from the University of California, Berkeley, Berkeley, CA, USA, in 1981 and the Ph.D. degree from the Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, in 1986.

He is currently a Scientist with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. He is a member of the NASA Science Team for the NASA ISRO Synthetic Aperture Radar (NISAR) mission planned for launch in 2021, and the Chair of the CEOS Working Group on Calibration and Validation—SAR subgroup. His research interests include studying wetland inundation extent with synthetic aperture radar.