This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

The Third Generation of Pan-Canadian Wetland Map at 10m Resolution Using Multi-Source Earth Observation Data on Cloud Computing Platform

Masoud Mahdianpari, Brian Brisco, Jean Granger, Fariba Mohammadimanesh, Bahram Salehi, Saeid Homayouni, and Laura Bourgeau-Chavez

Abstract*— Development of the Canadian Wetland Inventory Map (CWIM) has thus far proceeded over two generations, reporting the extent and location of bog, fen, swamp, marsh, and water wetlands across the country with increasing accuracy. Each generation of this training inventory has improved the previous results by including additional reference wetland data and focusing on processing at the scale of ecozone, which represent ecologically distinct regions of Canada. The first and second generations attained relatively highly accurate results with an average approaching 86% though some over-estimated wetland extents, particularly of the swamp class. The current research represents a third refinement of the inventory map. It was designed to improve the overall accuracy and reduce wetlands overestimation by modifying test and train data and integrating additional environmental and remote sensing datasets, including countrywide coverage of L-band ALOS PALSAR-2, SRTM, and Arctic digital elevation model, nighttime light, temperature, and precipitation data. Using a random forest classification within Google Earth Engine, the average overall accuracy obtained for the CWIM3 is 90.53%, an improvement of 4.77% over previous results. All ecozones experienced an overall accuracy increase of 2% or greater and individual ecozone overall accuracy results range between 94% at the highest to 84% at the lowest. Visual inspection of the classification products demonstrates a reduction of wetland area over-estimation compared to previous inventory generations. In this study, several classification scenarios were defined to assess the effect of preprocessing and the benefits of incorporating multi-source data for large-scale wetland mapping. In addition, the development of a confidence map helps visualize where current results are most and least reliable given the amount of wetland test and train data and the extent of recent landscape disturbance (fire). The resulting overall accuracies and wetland areal extent reveal the importance of multi-source data and adequate test and train data for wetland classification at a countrywide scale.

***Index Terms*— Random Forest, remote sensing, Multi-source data, Google Earth Engine.**

I. Introduction

UNTIL recently, the production of large-scale land cover maps through the classification of remote sensing observations required substantial amounts of time, labor, and complex methodologies. Additionally, the resolution of these maps tended to be coarse due to the nature of historically free remote sensing data such as MODIS (250 m) and Landsat (30 m) [1]. Despite such difficulties and limitations, large-scale land cover data are essential for a broad range of applications related to environmental management, climate change, and the assessment of major habitats. Examples of such land cover data in Canada include the 30m Annual Crop Inventory (ACI) [2], and the 30m Land Cover of Canada (LCC) [3], the former spanning the agricultural lands of southern Canada while the latter spanning the entire country [4]. These datasets provide crucial spatial information related to the location of numerous anthropogenic and non-anthropogenic land covers, including urban, agriculture, forest, herbaceous, and barren landscapes [5]. However, these datasets lack detailed wetland spatial information at the level of class. Such information that would be helpful for a multitude of environmental applications, given the different functions and distribution of wetlands at the class level [6]. An estimated 16% of Canada is currently covered in wetlands [7], and given the relatively recent and growing impacts of climate change (permafrost melt, changes to temperature and precipitation), wetland spatial data at the level of wetland class is an increasing necessity [8].

Masoud Mahdianpari is with the C-CORE, St. John's, NL A1B 3X5, Canada, and also with the Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada (e-mail: m.mahdianpari@mun.ca).

Brian Brisco is with The Canada Centre for Mapping and Earth Observation, Ottawa K1S 5K2, Canada (e-mail: brian.brisco@canada.ca).

Jean Granger and Fariba Mohammadimanesh are with the C-CORE, St. John's, NL A1B 3X5, Canada (e-mail: jean.granger@c-core.ca; fariba.mohammadimanesh@c-core.ca).

Bahram Salehi is with Environmental Resources Engineering, College of Environmental Science and Forestry, State University of New York, NY 13210, USA (e-mail: bsalehi@esf.edu).

Saeid Homayouni is with the Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, Quebec City, QC G1K 9A9, Canada (e-mail: saeid.homayouni@ete.inrs.ca).

Laura Bourgeau-Chavez is with Michigan Tech Research Institute, Michigan Technological University, Ann Arbor, MI, USA (e-mail: lchavez@mtu.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Wetlands are habitats characterized by a dominance of hydrophytic vegetation and saturated soils, though these characteristics manifest in various visually and ecologically distinct ways, which are sometimes grouped into different classes [9], [10]. In Canada, wetland classes can be defined following the Canadian Wetland Classification System (CWCS) [11]. The CWCS outlines five wetland classes of bog, fen, swamp, marsh, and shallow and open water based on shared broad vegetation and hydrological patterns. To briefly summarize the CWCS, bog wetlands are ombrotrophic peatlands dominated by sphagnum moss, fen wetlands are also peatlands, but are minerotrophic dominated by both moss and graminoid vegetation, swamp wetlands are dominated by woody vegetation, and marsh are wetlands that experience water fluctuations and are dominated by emergent herbaceous vegetation [12], [13]. Each class functions somewhat differently and in ways that benefit humans and other animals across the country and the globe via habitat provision, carbon storage, flood mitigation, and food provision, amongst many other benefits [14]. These five classes form the basis of wetland classification in Canada using remote sensing, but the products and methods are almost always implemented at small (at least relative to entire provinces and ecozones), geographical scales, such as that of watersheds, conservation areas, protected park, wildlife areas, municipalities, and at the scale of agricultural or industrial development [1].

The lack of large geographical scale wetland-class spatial information is likely the result of several factors, including limited wetland-related ground-truthing fieldwork, associated difficulties related to collecting wetland-related test and train data, difficulties inherent to the discrimination of wetland classes using remote sensing techniques, including the lower resolution of free Landsat data, and ecological characteristics inherent to wetlands [15]. For example, wetlands of different classes will often share visually similar vegetation patterns (such as bog and nutrient-poor fen) and are typically differentiated using field-validation of indicator species, nutrient quality, or sub-surface hydrology [11] all of which is not easily resolved by open remote sensing data [16]. Additionally, some wetland classes, such as marsh, experience dynamic changes to vegetation and hydrology over different seasons and are impacted by weather events such as rain, impacting spectral signatures captured by remote sensing data over time [17]. To make matters more difficult, most wetlands within close distances of roads and easily accessible locations have been damaged or destroyed. As such, acquiring wetland ground-truth data requires labor-intensive field campaigns. For all of these reasons, remote sensing of wetlands is a relatively challenging problem even at small (less than that of a province or ecozone) geographical scales [18].

In more recent times, however, there has been increased interest in wetland-class mapping [1]. This has resulted in a relatively substantial amount of research dedicated to mapping wetland classes around the world [19]. Additionally, there has been a boom in the production of large-scale remote-sensing thematic datasets, attributed to recent advancements in computational and software development, including cloud computing, and an increase in the amount and availability of multi-sensor remote sensing data sets. This boom has similarly resulted in more large-scale wetland thematic data. In China, for example, [20] produced a national-scale wetland map at the class level using object-based image analysis, hierarchical classification, and Landsat-8 imagery, estimating roughly 451,0484 km^2 of wetlands, a dominance of inland marsh, and rarity of coastal swamp wetlands. Similarly, in Canada, [6] assessed the status of wetlands at the level of treed wetland and non-treed wetland across forested ecozones of Canada over 33 years using Landsat imagery composites.

To address the data gap in Canada related to large-scale wetland spatial information at the class level, [21] developed the Canadian Wetland Inventory Map (CWIM), a product that describes wetland class across all of Canada using advanced remote sensing and cloud computing techniques. This project has been implemented over several generations, each improving on the last. The original CWIM (herein CWIM1) produced a 10m wetland inventory map of Canada using multi-year and multi-source (Sentinel-1 (S1) and Sentinel-2 (S2)) remote sensing data and an object-based random forest (RF) methodology within Google Earth Engine (GEE) [21]. Given the distribution of testing and training data available to the project at the time, provincial boundaries were selected as processing units. Overall accuracies (OA) ranged from 74% to 84%, depending on the province.

To improve on the results of the CWIM1, soon after, the second generation of the CWIM (herein CWIM2) was developed [22]. Changes to the original CWIM1 methodology included integrating a larger pool of wetland testing and training data, including filling some data gaps in Northern Canada and processing at the scale of ecozone rather than province. An ecozone-scale processing unit was chosen rather than a provincial-scale, given a greater geographical distribution of test and train data available to the CWIM2 and the ecologically relevant scale of ecozone units. Ecozones divide Canada into 15 ecologically distinct areas and are a more meaningful unit ecologically than political boundaries [23]. OA results ranged from 76% to 91%, a 7% improvement over the CWIM2. Despite the improvement, issues remained with an over-estimation of wetland classes, particularly swamp and lower accuracies in regions with little ground truth.

The purpose of this study is the implementation of the third generation of the CWIM (CWIM3), which will integrate more remote sensing datasets to improve OA and reduce wetland area over-estimation. Wetland-remote sensing research over the past 40 years [1] has demonstrated the value of multi-sensor and multi-feature methods to discriminate wetland classes better. Generally, in wetland-remote sensing research, higher OA and better class discrimination are achieved when integrating multiple features from multiple optical, multiple SAR, and various other datasets such as elevation, temperature, etc [8]. Such a multi-feature methodology is challenging to implement at a large-scale given restriction in data coverage of some remote sensing datasets (e.g., the Canadian DEM is not

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

present in Northern Canada), cost (LiDAR and other higher spatial resolution data across Canada are limited and costly to obtain), and difficultly as a result of computation power and processing. However, with time and through collaboration, advances in the technical capabilities to integrate multiple datasets for large-scale classification are becoming more feasible to be taken advantage of by the CWIM3.

As such, this research aims to develop the third generation of the CWIM, which will be developed by integrating a multitude of new datasets to improve wetland class discrimination. These datasets include ALOS PALSAR-2, 10m Canada-wide elevation data, city light information, and climate data (temperature and precipitation). Additional effort has been dedicated to refining the test and train datasets within each ecozone across Canada. Specific objectives are to (1) improve the accuracy of the CWIM3 compared to the CWIM2, (2) reduce wetland class area overestimation, and (3) improve on the processing time required to produce a classified wetland map for each ecozone. Several research questions are also answered by defining different classification scenarios, which determine the effect of preprocessing steps, integration of various sources of remote sensing and non-remote sensing data, and processing units (i.e., ecozone-by-ecozone vs. the entire country) on wetland classification accuracy. The results are then compared to other similar large-scale Canadian classification datasets.

II. Study Area

The study area encompasses the entire landmass of the country of Canada, totaling 9.9 million km^2. Processing was implemented at the scale of ecozone. Canada is divided into 15 ecozones, the boundaries of which define an ecologically distinct area characterized by interacting biotic and abiotic factors [23]. Ecozones often cross multiple provincial boundaries and range in size between 117,240 km^2 at the smallest to 1,857,530 km^2 at the largest. Table I in [22] summarizes the general landscape characteristics of each ecozone. For purposes of this research, we modified some ecozone boundaries due to limited testing and training data distribution, leaving 13 ecozone processing units. As was implemented in the CWIM2, we group the three ecozones that comprise Northern Canada (Southern Arctic, Northern Arctic, and the Arctic Cordillera) due to the limited amount of wetland test and train data available in this part of Canada. This area is referred to as the Northern Ecozones herein. For similar reasons we group the Boreal and Taiga Cordillera into a single unit, named Boreal/Taiga Cordillera ecozone. Given the size and abundance of training data in the Boreal Shield ecozone, we split the Boreal Shield down the middle into the Boreal Shield West and Boreal Shield East for ease of processing. See Fig.1 for the distribution of these ecozones across Canada.

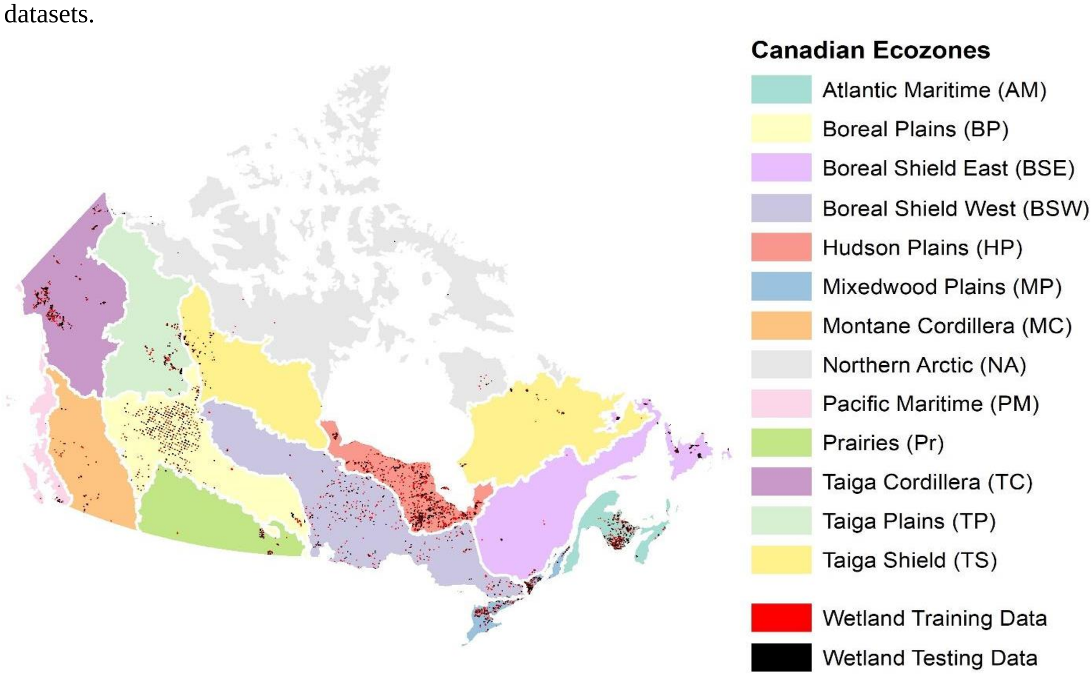


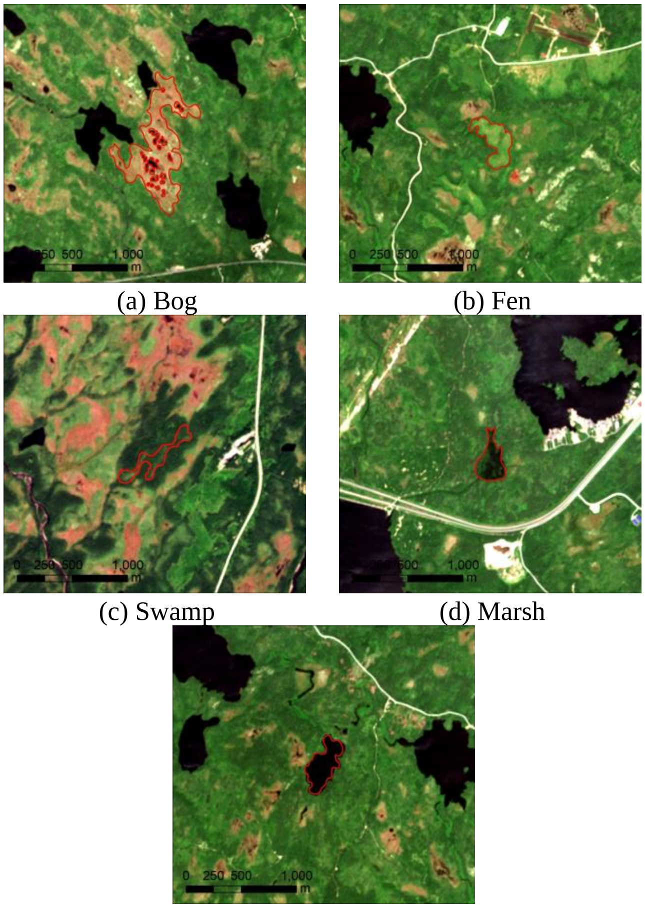
Fig. 1. Canadian Ecozones, modified for purposes of implementing the CWIM3. Wetland testing and training data are visible in black and red.

III. Methods

A. Test and Train Data Preparation

Wetland test and train data (the distribution of which can be seen in Fig. 1) has been sourced from many partners to produce the CWIM. Because these datasets were collected under varying circumstances and for differing purposes, an effort was made to better standardize and improve the cohesiveness of these wetland datasets before producing the CWIM2 [22]. This included modifying wetland boundaries, altering class labels, removing potentially inaccurate polygons, and filtering by size by removing any polygons smaller than one hectare and greater than 100 hectares because small polygons would not contain any helpful spectral information for the classifier and large polygons had a higher chance of being highly spectrally heterogeneous [22]. A sample of the testing and training data polygons can be seen in Fig. 2.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing



(a) Bog (b) Fen (c) Swamp (d) Marsh (e) Water

Fig. 2. Examples of wetland polygons that comprise the testing and training datasets used in developing the CWIM3 overlaid Sentinel-2 summer imagery.

To help improve the results of the CWIM3, additional effort was dedicated to improving the quality and quantity of the non-wetland testing and training data. Non-wetland data helps to reduce over-classification of wetland areas in remote sensing supervised classification methods. For example, a dataset with a representative sample of forest data can help to reduce over-classification of woody wetlands, such as swamp. An issue with the test and train data applied to the CWIM2 was an excess of wetland test and train data relative to non-wetland test and train data. This likely contributed to an over-estimation of wetland classes in certain ecozones, particularly swamp. [24] suggest that a quality testing and training dataset represents the general land cover of the study area. As such, the ratio of wetland and non-wetland data was modified to ensure a more considerable amount of non-wetland polygons in most ecozones. The Hudson Plains ecozone is an exception given its overwhelming dominance by wetlands. Because there was a limited amount of non-wetland land cover data provided directly to this project, non-wetland data was obtained via governmental datasets such as the 2015 Land Cover [3]. Considered upland classes included forest, shrubland, grassland, agriculture, urban, and barren/exposed, though these are reported as a single land cover class (non-wetland) in the final results.

TABLE I

Wetland and non-wetland test and train polygons per ecozone. Data in bold text highlights ecozones with low amounts of wetland test and train data relative to other ecozones. Ecozone abbreviations are as follows: Atlantic Maritime (AM), Boreal and Taiga Cordillera (BCTC), Boreal Plains (BP), Boreal Shield East (BSE) and West (BSW), Hudson Plains (HP), Mixedwood Plains (MP), Montane Cordillera (MC), Northern Ecozones (NE), Pacific Maritime (PM), Prairies (Pr), Taiga Plains (TP), and Taiga Shield (TS).

Land Cover Classes

	Bog		Fen		Swamp		Marsh		Water		Non-wetland	
Ecozones	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train
AM	71	30	182	76	163	76	132	45	133	57	1009	425
BCTC	103	44	92	39	149	64	92	41	148	63	1619	696
BP	133	56	378	163	108	46	133	59	119	51	1037	442
BSE	216	99	232	100	167	70	108	53	72	33	923	377
BSW	118	49	126	54	99	41	73	33	47	19	638	267
HP	438	185	392	170	130	55	69	30	56	24	334	144
MP	68	31	149	64	242	104	130	60	42	18	971	420
MC	**na**	**na**	**11**	**5**	**25**	**11**	**27**	**9**	**18**	**7**	350	152
NE	**na**	**na**	**42**	**20**	**63**	**26**	**23**	**12**	**91**	**35**	2241	977
PM	**16**	**6**	**31**	**14**	**23**	**10**	**46**	**19**	**20**	**9**	314	144
Pr	**na**	**na**	**29**	**11**	**41**	**15**	**43**	**19**	**69**	**29**	521	227
TP	**43**	**19**	**97**	**39**	**21**	**9**	**44**	**20**	**43**	**18**	542	240
TS	**53**	**20**	**71**	**31**	**55**	**24**	**65**	**27**	**90**	**38**	477	204

The final test and train datasets used to produce the classification for each ecozone are outlined in Table I. In total, the final dataset is comprised of 8804 wetlands and 15691 non-wetland polygons. In each ecozone, the dataset was split 70/30 into training and testing datasets, respectively. Note that due to the limited amount of wetland data available in some ecozones, the bog class was not considered in the Northern Ecozones, Montane Cordillera, and Prairies. However, any occurrence of bog in these ecozones will likely be classified as fen. Bog and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

fen share many similar ecological features, and it was deemed acceptable to consider only the fen class.

B. Satellite Imagery Processing

All satellite imagery was processed in the GEE cloud computing platform [25]. In this study, the GEE data catalog was employed to collect satellite imagery over different Canadian ecozones during the summers of 2017-2020 from S1 and S2 and 2017-2018 to develop an ALOS PALSAR-2 yearly mosaic.

The S1 mission provides data from a dual-polarization C-band Synthetic Aperture Radar (SAR) instrument. This collection includes the S1 Ground Range Detected (GRD) scenes, processed using the S1 Toolbox to generate a calibrated, ortho-corrected product [26]. The collection is updated daily. New assets are ingested to GEE within two days after they become available. In this study, a total of 6,222 and 27,102 Level-1 S1 GRD images were acquired in the HH-HV and VV-VH polarization modes, respectively. Different preprocessing steps, including thermal noise removal, radiometric calibration, terrain correction using SRTM 30 (or ASTER DEM for areas greater than 60 degrees latitude, where SRTM is not available) were carried out in GEE on each scene of S1 data (i.e., interferometric wide (IW) mode with a resolution of 10m). An adaptive Lee sigma filter with a pixel size of 7x7 was then applied To reduce the speckle noise from S1 data. Median and standard deviation mosaics of the time stacks of S1 imagery were then extracted and employed for wetland classification.

S2 is a wide-swath, high-resolution, multispectral imaging mission with a global five-day revisit time. The S2 Multispectral Instrument (MSI) collects data in 13 spectral bands: visible and NIR at 10 meters, red edge, and SWIR at 20 meters, and atmospheric bands at 60 meters spatial resolution. In this study, S2 Surface Reflectance (SR, Level-2A) and Top of Atmosphere (TOA, Level-1C) imagery were collected on a tri-monthly period, from June to August. This is because generating a 10-m cloud-free Sentinel-2 composite for Canada over a shorter time was challenging. A total of 115,747 Sentinel-2 images (with less than 20% cloud cover) from summer 2017 to 2020 were queried from the GEE data catalog. Novel to the CWIM methodology is an L-band ALOS PALSAR-2 mosaic, a seamless SAR dataset created by mosaicking ALOS PALSAR-2 SAR imagery strips. In this dataset, the strip data were selected through visual inspection of the browse mosaics available over the period, with those showing minimum response to surface moisture preferentially used. Several optical and SAR features were extracted from these satellite imagery and were incorporated into the classification step (see Table II and Fig. A).

TABLE II

A list of extracted bands, features, indices, and auxiliary data used in this study.

Type	Source	Spatial Resolution	Time-scale	Parameters
Spectral features	Sentinel-2 (band reflectance)	10m	2017-2020	Blue, Green, Red, Red Edge1, Red Edge 2, Red Edge 3, NIR, Red Edge 4, SWIR 1, and SWIR 2
	Sentinel-2 (Spectral indices)	10m	2017-2020	Normalized Difference Water Index (NDWI), Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI), Ratio Vegetation Index (RVI), Normalized Difference Built-up Index (NDBI), Normalized Burn Ratio (NBR), Normalized Difference Snow Index (NDSI), and Bare Soil Index (BSI).
SAR features	Sentinel-1 (HH+HV)	10m*	2017-2020	HH polarization backscattering coefficient, HV polarization backscattering coefficient, Span, Ratio
	Sentinel-1 (VV+VH)	10m*	2017-2020	VV polarization backscattering coefficient, VH polarization backscattering coefficient, Span, Ratio
	ALOS PALSAR-2	25m*	2017-2018	HH polarization backscattering coefficient, HV polarization backscattering coefficient, Span, Ratio
Topographical features	SRTM DEM	10m		Elevation, Slope, Aspect
	ArcticDEM	10m		Elevation, Slope, Aspect

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Environmental Features	ERA5	25km	2010-2020 (year average)	Temperature, Precipitation
Nighttime lights	Visible Infrared Imaging Radiometer Suite (VIIRS)	10m	2017-2020	Day/Night Band (DNB)

*These numbers represent the SAR pixel spacing.

C. Auxiliary Data Preparation and Processing

The use of exclusively spectral classification models for large-scale land cover mapping may suffer due to dramatic changes in climatic and ecological characteristics across geographical gradients such as ecozones and affect the final classification accuracy. For example, the ecological characteristics of wetland classes such as vegetation composition and structure, soil type, and hydrology can vary under the ecozones' climatic and ecological parameters. Thus, a fen presented in the Atlantic Maritime and characterized by a maritime climate (i.e., cool and moist) may appear spectrally different from a fen in the Montane Cordillera ecozone, with mainly continental climate (warm and dry). Thus, signatures of wetland classes illustrate a possibly wide range of species composition, vegetation physiognomy, and land management strategy, all of which are combined to represent a single land cover class over a large geographic region in the final classification product.

To address such problems associated with large-scale land cover mapping, two common strategies have been employed in the literature: (a) dividing large-scale study areas into several small parts and applying classification models within small regions [27] and/or (b) incorporating environmental indices into the classification scheme, which allow classification models to take into account regional variations within different ecozones [28] . Both techniques have been examined in this study to identify which method is more effective for improving wetland classification results.

A Canada-wide digital elevation model was introduced to the CWIM to improve wetland discrimination. In particular, it is expected that adding DEM data will improve swamp class discrimination and help to reduce wetland area overestimation. The 30 m SRTM data covering southern Canada was resampled to 10 m and used alongside the 10 m Arctic DEM covering Northern Canada. Slope and aspect were also extracted from DEM and added to the classification scheme. Nighttime light data was used as another input feature. The Defense Meteorological Program (DMSP) Operational Line-Scan System (OLS) has a unique capability to detect visible and near-infrared (VNIR) emission sources at night. In particular, the nightlight data is a monthly average radiance composite image using nighttime data from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB). This dataset helps distinguish artificial surfaces from other land covers. Finally, climate data, including temperature and precipitation, with a resolution of 25 km, were added to our analysis from the ERA5 fifth-generation ECMWF atmospheric reanalysis of the global climate (Copernicus Climate Change Service, 2017). In particular, a 10-year (2010-2020) average and standard deviation in monthly precipitation and temperature were extracted from the climate data. It is expected that long-term precipitation data capture spectral differences between wetland classes in ecozones with different climates, such as the Atlantic Maritime (maritime climate), the Montane Cordillera (continental climate), and parts of the Northern Ecozones (the coldest and driest of all ecozones). Long-term temperature data also helps to capture the timing of maximum vegetation growth within different ecozones. All features extracted from satellite imagery and auxiliary data were then incorporated into an object-based classification scheme in various classification scenarios. A visual illustration of auxiliary datasets can be found in the Appendix.

D. Classification and accuracy assessment

In this study, an object-based classification scheme consisting of a simple non-iterative clustering (SNIC) method and an RF algorithm were used. Object-based classification was chosen as it produces objects that are more meaningful and tends to produce higher overall accuracies when classifying wetlands compared to pixel-based classification [1]. SNIC is a non-iterative, region-growing approach for generating superpixels, wherein centroids of clusters are evolved based on online averaging. SNIC uses a priority queue, 4- or 8-connected candidate pixels to the currently growing superpixels cluster and gives a higher priority to the pixels with the smallest distance from the centroid to join the cluster [29]. The algorithm takes advantage of both priority queue and online averaging to evolve the centroid once each new pixel is added to the given cluster. Accordingly, SNIC is faster and demands less memory relative to similar clustering algorithms (e.g., Simple Linear Iterative Clustering). This is attributed to the introduction of connectivity (e.g., 4- or 8-connected pixels) from the beginning of the algorithm, resulting in fewer distances during centroid evolution.

The RF algorithm was implemented for classification in this study. RF is an ensemble learning method comprised of a group

TABLE III

Wetland classification scenarios examined in this study.

Scenario	Wetland classification scenarios	Features	Objective	Classification scale strategy
1	Optical top-of-atmosphere data	Features extracted from Sentinel-2 TOA	The effect of preprocessing	Classification applied ecozone-by-ecozone

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

2	Optical surface reflectance data	Features extracted from Sentinel-2 SR	The effect of preprocessing, the usefulness of optical data	Classification applied ecozone-by-ecozone
3	SAR data	Features extracted from Sentinel-1 and ALOS data	The usefulness of SAR data	Classification applied ecozone-by-ecozone
4	Optical and SAR	Features extracted from Sentinel-1, ALOS, and Sentinel-2 SR data	The importance of combining optical and SAR data	Classification applied ecozone-by-ecozone
5	Optical, SAR, and DEM	Features extracted from Sentinel-1, ALOS, and Sentinel-2 SR data along with DEM	The importance of applying DEM and the classification scale	Classification applied ecozone-by-ecozone
6	Optical, SAR, DEM, and environmental data	Features extracted from Sentinel-1, ALOS, and Sentinel-2 SR data, DEM, precipitation, temperature, and nighttime data	The effect of applying auxiliary data	Classification applied ecozone-by-ecozone
7	Optical, SAR, and DEM	Features extracted from Sentinel-1, ALOS, and Sentinel-2 SR data, DEM	The classification scale	Classification applied to the entire country
8	Optical, SAR, DEM, and environmental data	Features extracted from Sentinel-1, ALOS, and Sentinel-2 SR data, DEM, precipitation, temperature, and nighttime data	The effect of applying auxiliary data and the classification scale	Classification applied to the entire country

of tree classifiers handling high-dimension remote sensing data [30]. As such, RF is not prone to overfitting and performs well with noisy input data. Assigning a label to each object is based on the majority vote of trees [31]. RF can be tuned by adjusting two input parameters, namely the number of trees (Ntree), which is generated by randomly selecting samples from the training data, and the number of variables (Mtry) used for tree node splitting. An automated hyperparameter tuning was employed to select the Ntree of 100 and Mtry set to the square root of the number of features [32].

In this study, several classification scenarios were defined to assess the effect of preprocessing and the benefits of incorporating multi-source data for large-scale wetland mapping (see Table III). In particular, both TOA and SR S2 data were used to identify the importance of applying atmospheric correction on the final classification results. Next, three classification scenarios were defined to determine the usefulness of combining optical and SAR data for large-scale wetland applications [10]. The effect of adding auxiliary data, including DEM, environmental data (i.e., precipitation and temperature), and nighttime data were also explored. The importance of applying classification models at various scales (i.e., ecozone-by-ecozone vs the entire country) was determined by comparing classification models applied ecozone-by-ecozone versus the entire country. This also identifies whether auxiliary data are more influential to the classification results or applying classification models in small ecozones.

Overall accuracy (OA) and Kappa coefficients were used to evaluate the capability of the wetland classification in each ecozone. In addition, the average F1-score for wetland and non-wetland classes were measured. F1-score (range 0–1) is the harmonic average of precision and recall and is useful for unbalanced validation data.

IV. RESULTS

Fig. 3 compares the classification accuracies achieved under different wetland classification scenarios outlined in Table III. Comparing classification scenarios 1 and 2 reveals that atmospheric correction of S2 data is essential, as an improvement of about 2.5% was achieved when surface reflectance data is used. Overall, the classification accuracy obtained from single source SAR data is significantly lower than single-source optical data for wetland mapping (see S2 vs. S3). However, the inclusion of both types of data (i.e., optical and SAR) improved the classification accuracy by about 6% compared to single-source optical data (see S2 and S4) and 18% relative to exclusive use of SAR data (see S3 vs S4). An additional 2% improvement is obtained through the inclusion of DEM data. This is attributed in part to the improvement in discrimination between the forest and swamp classes. Finally, the classification accuracy exceeded 90% when other auxiliary data, namely precipitation, temperature, and nighttime data, were incorporated in the classification scheme in scenario 6.

Regarding the classification scale for large-scale applications, the results confirmed the necessity for performing

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

TABLE IV

OA results for the CWIM3 and the CWIM2 for each ecozone.

		CWIM3				CWIM2		Change	
		F1-score Wetland	*F1-score Non-wetland*	*OA(%)*	*Kappa*	*OA(%)*	*Kappa*	*OA(%)*	*Kappa*
Ecozone	*AM*	0.83	0.97	94	0.9	88	0.87	6	0.03
	BCTC	0.69	0.93	84	0.8	76	0.73	8	0.07
	BP	0.78	0.95	89	0.87	87	0.86	2	0.01
	BSE	0.79	0.94	91	0.87	86	0.84	5	0.03
	BSW	0.81	0.96	92	0.89	87	0.86	5	0.03
	HP	0.83	0.96	93	0.9	88	0.87	5	0.03
	MP	0.82	0.94	93	0.89	88	0.87	5	0.02
	MC	0.76	0.95	88	0.85	85	0.83	3	0.02
	NE	0.77	0.96	92	0.87	89	0.87	3	0
	PM	0.80	0.94	89	0.86	84	0.82	5	0.04
	Pr	0.84	0.97	94	0.91	91	0.9	3	0.01
	TP	0.75	0.93	86	0.84	82	0.79	4	0.05
	TS	0.81	0.96	92	0.89	84	0.79	8	0.1

classification models ecozone-by-ecozone rather than the entire country. For example, classification scenarios 5 and 7, as well as 6 and 8 use the same input features albeit within different geographic scales. In both cases, significant improvement was achieved for classifications through the ecozone-by-ecozone strategy. Although previous studies suggested that for large-scale land cover mapping either inclusion of auxiliary data or applying different classification models within a small area should be sufficient [33]. This does not hold for a country like Canada, where ecological and climatic features can vary even within a single ecozone.

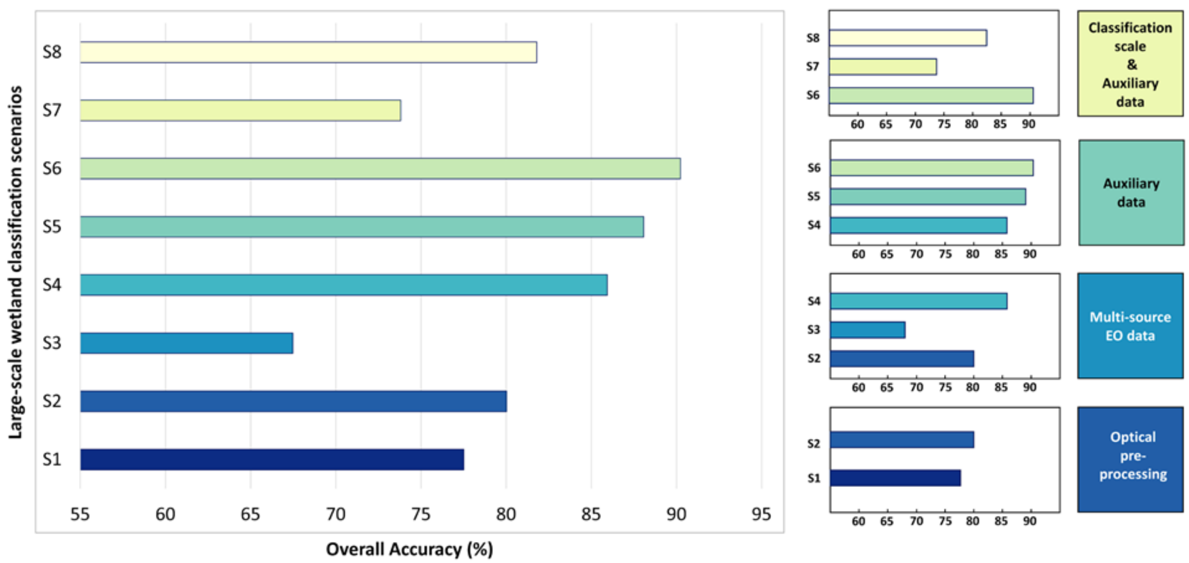


Fig. 3. Comparisons between classification accuracies obtained for different classification scenarios outlined in Table III.

Based on the results of classification scenarios, classification scenario S6 was selected to produce final classification results. S6 produced an average overall accuracy of 90.53% and an average Kappa coefficient of 0.87 across all ecozones. This is an increase of 4.77% in terms of OA compared to CWIM2, which has an average OA of 85.76%. At the level of ecozone, OAs range between 94% at the highest (Atlantic Maritimes and Prairies) and 84% at the lowest (Boreal and Taiga Cordillera). This pattern is similarly reflected in the CWIM2, though with lower OA percentages, the highest being 91% (Prairies) and the lowest being 76% (Boreal and Taiga Cordillera). Table IV further outlines the OA percentages across each ecozone for both the CWIM2 and the CWIM3. For each ecozone, the OA percentage increased by at least 2% and at most 8%. The smallest OA increase compared to the CWIM3 occurred in the Boreal Plains ecozone, at 2%. The greatest OA increase compared to the CWIM2 occurred in the Taiga Shield ecozone at 8%. A majority of ecozones (5 out of 13) experienced an OA increase of 5%.

Across Canada, the results of the CWIM3 reveal an estimated 16.69% of wetlands. Fig. 4 displays the distribution of wetlands across the country for each class. Spatial patterns of wetland classes are well preserved in the map, and the prevalence of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

wetland classes is clear in the Hudson Plains and Boreal Plains ecozones. Wetlands generally follow a central longitudinal distribution across the country and are less common in the north and the south. Compared to other areas, the water class is less prevalent on the west side of the country.

Table V outlines the percentage of wetlands per ecozone according to the CWIM3. The results of the CWIM3 are compared with other estimates of Canadian wetland coverage per ecozone by other related research investigating wetland change detection across Canada's forested ecozones [6] and estimates of wetlands extent by Environment and Climate Change Canada [7]. The Hudson Plains ecozone has the greatest total area of wetland, followed by the Boreal Plains and Taiga Plains. Ecozones with the fewer areas of wetlands are the Boreal and Taiga Cordillera ecozone and the Montane Cordillera. This pattern is similarly reflected in [3] and [9].

Fig. 7 displays the class composition of the total wetland area in each of the 13 ecozones. Excluding water, fen and bog are the most dominant wetland classes in Canada, followed by swamp and marsh. The dominant wetland class (excluding water) in the Hudson Plains ecozone is bog, followed by fen, then swamp, and marsh. The dominant wetland class in the Boreal Plains ecozone is fen, followed by bog, then marsh, and swamp. In the Boreal and Taiga Cordillera ecozones, fen and swamp cover the greatest wetland areas, followed by bog and marsh. In the Montane Cordillera ecozone, the marsh covers most of the areas, followed by fen and swamp. Visual comparison of ecozones across the CWIM2 and CWIM3 reveals a substantial reduction in the total amount of overestimated wetland area in all ecozones. Visual analysis of the Boreal Plains ecozone, for example, seen in Fig. 5, reveals a reduction of total wetland areas and a better concentration of those wetlands along the northern region of the ecozone when comparing the CWIM2 to the CWIM3 results. Classification noise is also reduced between CWIM generations.

Table V
Percent coverage of wetlands per ecozone as reported by the CWIM3, [7] and [6].

Ecozone	**CWIM3 (%)**	[7] **(%)**	[6] **(%)**
AM	9.06	6.30	13.94
BCTC	3.81	2.40	1.25
BP	27.30	30.30	14.06
BSE + BSW	20.83	16.90	15.81
HP	86.16	78.80	80.88
MP	9.29	11.10	NA
MC	4.19	1.86	0.19
NE	6.21	9.45	NA
PM	4.71	1.12	4.17
Pr	5.55	3.10	NA
TP	24.88	25.00	29.20
TS	15.00	11.00	13.53
Total	**16.69**	**16.00**	**16.95**

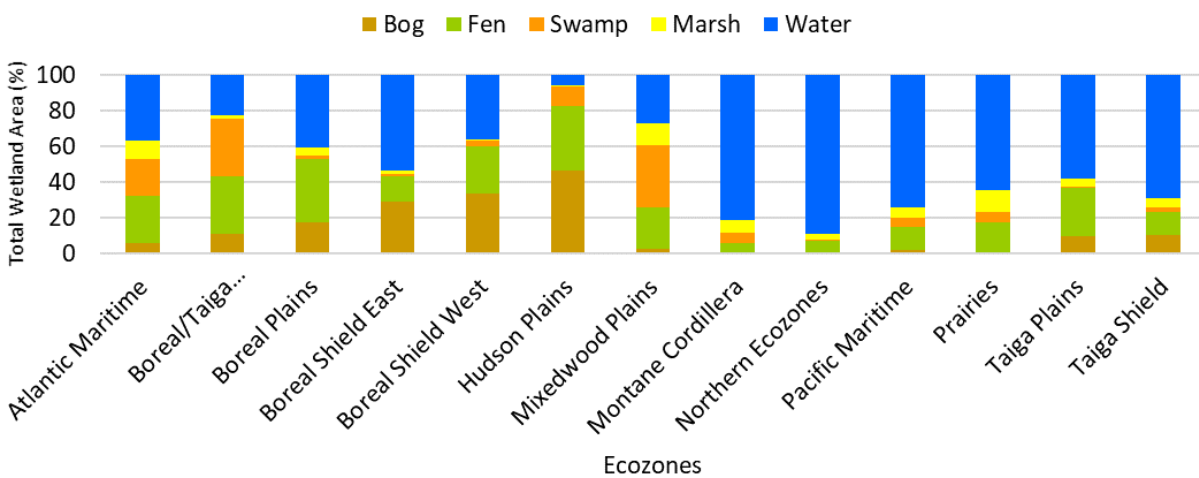


Fig. 4. The class composition of the total wetland area in each ecozone.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

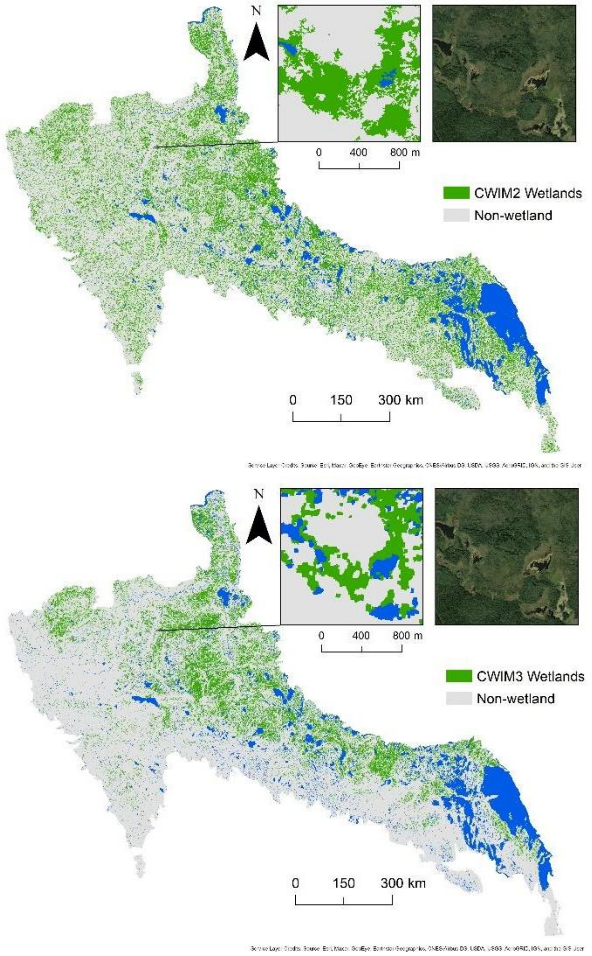


Fig. 5. Comparison of the Boreal Plains wetland classification results for the CWIM2 (top) and CWIM3 (bottom). Green represents non-water wetland, blue represents water, and grey represents non-wetlands.

Similar results can be seen in the Taiga Shield ecozone (Fig. 6), where wetland areas are better concentrated within areas characterized by lowlands and plains [34] along the ecozone boundary on the northeast side. Compared to the CWIM2, there is much less wetland area in the south and along the south-west boarder of the boundary that lays along the Mackenzie Mountain range. Again, there is a reduction in classification noise between map generations, resulting in a more clear visualization of areas that contain a high concentration of wetland area.

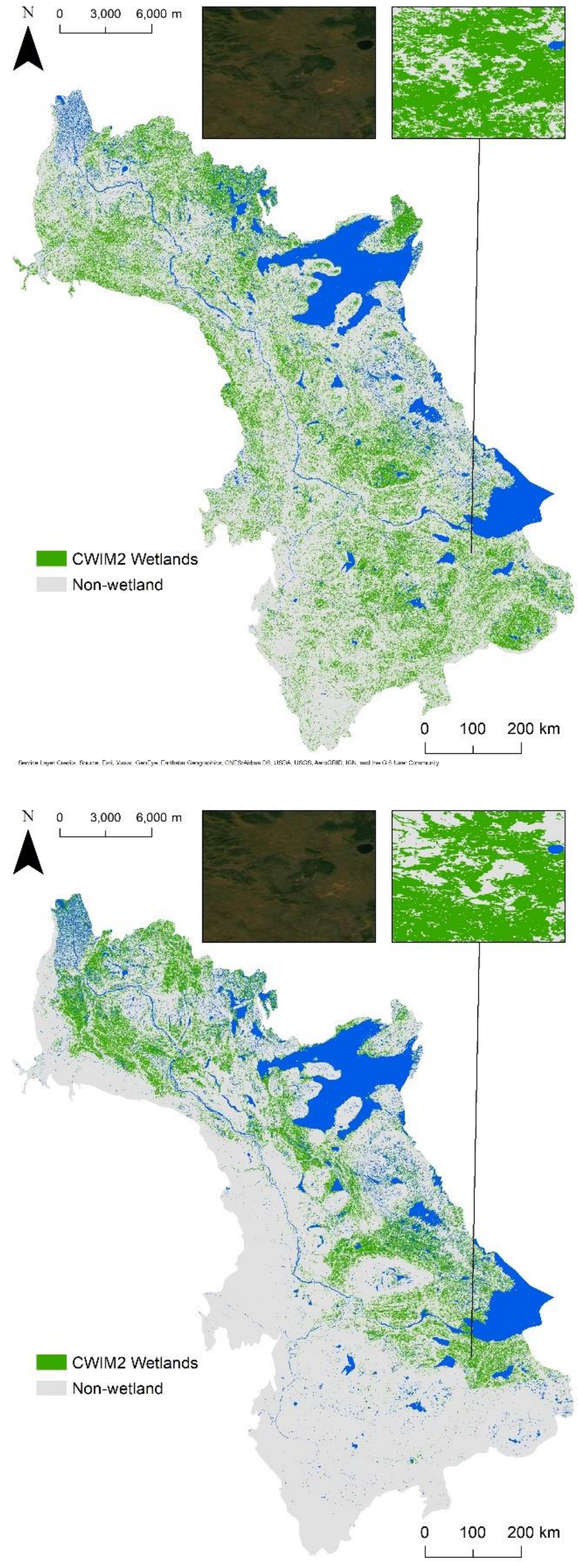


Fig. 6. Comparison of the Taiga Plains wetland classification results for the CWIM2 (top) and CWIM3 (bottom). Green represents wetland, blue represents water, and grey represents non-wetlands.

Fig. 7 shows the minimal wetland area present in the highly agricultural Prairies ecozone described by the CWIM2 and CWIM3. Though changes appear to be broadly minimal between these generations, wetlands that were not captured

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

previously in the CWIM2 have been captured by the CWIM3. This is particularly the case along the east side of the ecozone in and around Lake Winnipeg, where there is a higher concentration of wetlands (particularly peatlands) relative to the rest of the highly developed area. As is the case with other ecozones, general wetland noise has been reduced between generations as well.

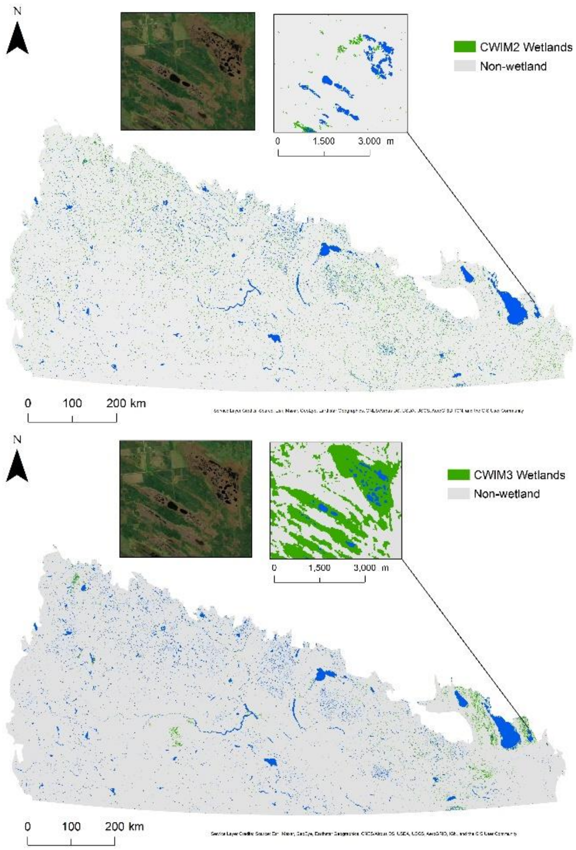


Fig. 7. Comparison of the Prairies wetland classification results for the CWIM2 (top) and CWIM3 (bottom). Green represents non-water wetland, blue represents water, and grey represents non-wetlands.

Fig. 8 also compares the classification of wetlands at a location in the Prairies ecozone from various data sources, including the CWIM3, CWIM2, 2015 LCC, and ACI maps. When compared to both the CWIM2 and the 2015 LCC, the CWIM3 better captures the extent of the wetlands as seen on the ground (in the optical imagery at the bottom of the figure), particularly those wetland areas that are long and thin in shape. Generally, the CWIM2 and the 2015 LCC datasets underestimate overall wetland area at this location in the Prairies ecozone. The CWIM3 results, at least in terms of wetland extent, is comparable to that seen in the ACI, however the CWIM3 provides the added benefit of discriminating wetland area at the level wetland class, rather than only describing wetlands as a single class, as is the case with the ACI dataset.

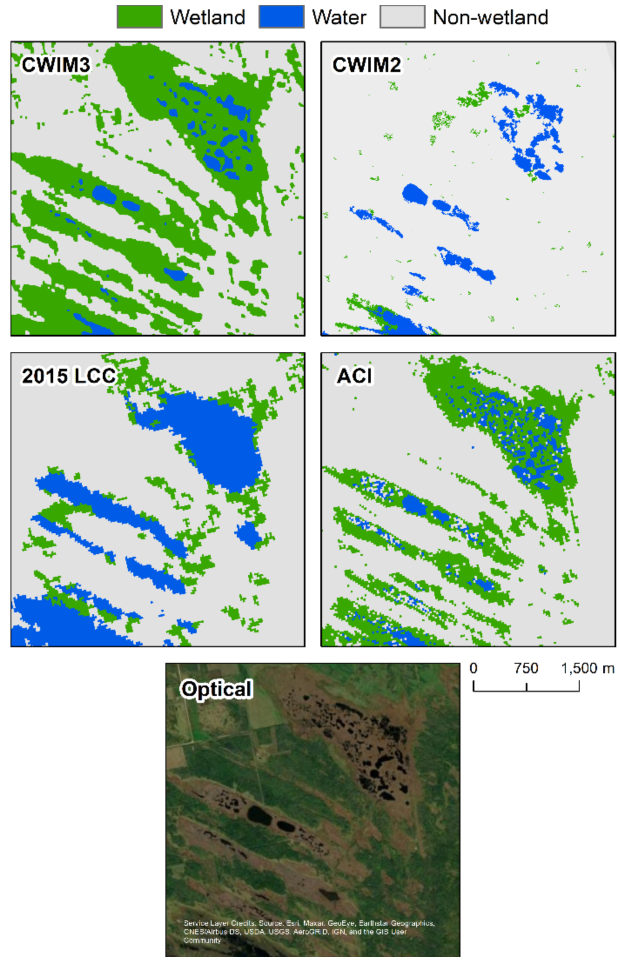


Fig. 8. Classification of wetland area in the Prairies ecozone from various data sources, including the CWIM3.

V. DISCUSSION

The resulting pan-Canadian wetland map here extends on our previous work focusing on generating high-resolution wetland data, by which an overall improvement of about 10% and 5% in accuracy obtained relative to the CWIM1 [21] and CWIM2 [22], respectively. The accuracy of 10m resolution CWIM3 produced here is 90.53% that are comparable with other Sentinel-based large-scale land cover mapping globally [33], [35]. However, general land cover classes (e.g., water, bareland, and cropland) are much easier to be delineated compared to ecologically similar wetland classes separated in this study. The results are also comparable with Landsat-based large scale wetland maps produced in China [20]. For example, a recent study focusing on national wetland mapping in China reported an accuracy of 95.1% using Landsat data. This, however, was obtained with four rounds of manual editing which improved the accuracy from 80.6% to 95.1% [20]. Although a direct comparison between the accuracy obtained from the pan-Canadian Sentinel-based wetland maps (i.e., the CWIM generations) with the Canada-wide Landsat-based map [6] and wetland maps from other sources [7] is impossible, as the accuracies have not been reported from the latter studies, there is a general agreement between areal percentages of wetlands found in this study with the existing literature.

Given the inherent difficulties associated with wetland classification using automated remote sensing methods [15] and the variation in the amount of wetland testing and training data available to the CWIM3, the accuracy of the CWIM3 will vary across space, and it is likely that in certain areas, the accuracy will be less than the stated OA. Additional confounding factors, such as natural disasters like fire, can also reduce the accuracy of wetlands in disturbed areas. To better communicate this issue, a confidence map was developed using testing and training data distribution and the area of recent fires from 2010 to 2020 [36]. The confidence map is displayed in Fig. 9 and was developed using a simple multi-criteria analysis. In this map, the darker colors represent areas with greater confidence in the results of the wetland map, whereas the lighter colors represent areas where there is less confidence in the results of the classified map. Generally, confidence decreases moving north as a result of a lack of substantial testing and training data.

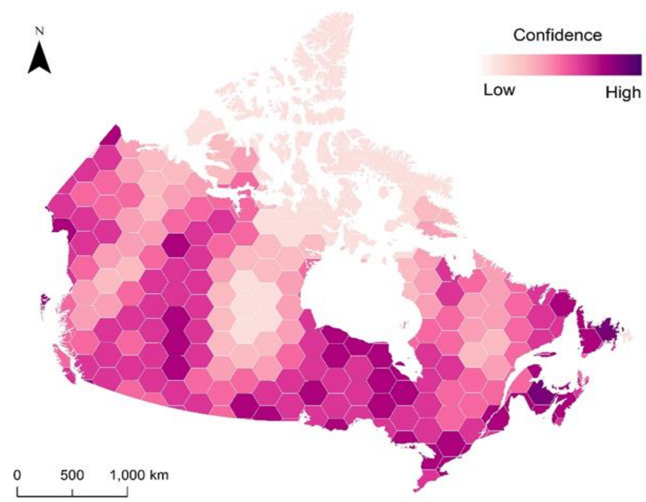


Fig. 9. Confidence in the accuracy of the CWIM based on wetland testing and training data distribution and location of recent fires.

Despite the substantial contribution of wetland training and testing data from many partners, there remain large expanses across many ecozones where there is little or no wetland data. Although [24] suggests that an optimal dataset is well distributed, this is a difficult challenge to address given the large size of most of these ecozones and given the substantial portion of Canada that is generally inaccessible to standard field campaigns. In Fig. 1, it can be seen that there is a lack of data in ecozones in northern parts of Canada, particularly in the Northern Ecozones and Taiga Shield given the large sizes of these ecozones. There is also a relative lack of testing and training data in the Pacific Maritime and Montane Cordillera ecozones in western Canada.

Consideration should also be given to the issue of spatial autocorrelation, particularly when considering the OA results per ecozone, and the reliability of the CWIM3 map. Spatial correlation could result in an overestimation of accuracy when it is not assessed. Because spatial autocorrelation of the testing and training data was not considered during this research (nor during the research to develop the CWIM2), and given that spatial autocorrelation is inherent to remote sensing data, the CWIM3 OA results are likely to be higher than what is represented by the map when compared to real life. Spatial autocorrelation will certainly contribute to OA results because our wetland datasets were generally collected via field campaigns that cover only very small geographical areas, and were collected not for non-remote sensing purposes. Consider the Taiga Plains ecozone, where wetland data available to this research is concentrated in a small area along the south-east, or the Boreal Shield East ecozone, where wetland data is largely concentrated in and around some populated areas in Newfoundland. Future generations of the CWIM should address or assess the issue of spatial autocorrelation, using Moran's I or implementing recent advances by [32].

Comprehensive classification of each ecozone is necessary to ensure wetlands are not overestimated, thus the need for non-wetland test and train data. For the CWIM3, this information was obtained from the 2015 LCC [2] dataset available via the Government of Canada. The accuracy of this dataset is variable across land cover classes and geographical areas. As such, some of the non-wetland land cover test and train data used in the CWIM3 are likely to include mixed land cover signatures. Future work may dedicate effort to improving the boundaries of these non-wetland land cover test and train data to include less land cover mixing, particularly along the polygon boundaries. This may also require refinement of the number of non-wetland land cover classes considered. The CWIM3 considered forest, shrubland, grassland, agriculture, urban, and barren land cover. However, an increase in the number of non-wetland classes considered (sub-grassland classes, sub-forest classes, sub-shrubland classes) may help to increase accuracy.

Improvements across the CWIM2 and CWIM3 is a result of many changes made in the processing and integration of spatial data across Canada, such as the inclusion of additional environmental datasets and satellite data. Changes to OA are also a result of direct modifications made to training and testing data inputs between the CWIM2 and CWIM3. The CWIM2 was developed using datasets that were wetland-dominant, regardless of actual proportion of wetland area in the landscape. However, based on research by [24], a choice was made to improve the relative proportions of non-wetland and wetland training and testing data, based on the general landscape of each ecozone, while developing the CWIM3. For example, most ecozones are not dominated by any wetland class, rather are dominated by forest. Thus, training and testing data in most ecozones were modified to include a greater proportion of forest training and testing data relative to wetland. As such, improvements to OA across CWIM generations is not only due to integration of new environmental and satellite datasets, but is likely a result of changes to training and testing datasets. In this research, we do not assess how much of the change to OA in each ecozone is due to the modified training and testing data set, however it is likely not negligible.

The resolution of the CWIM3 should be taken into consideration, particularly when examining areas in and around developed areas. Wetlands in these areas tend to be fragmented, have modified vegetation patterns, and are often smaller in size beyond the resolving power of 10m satellite data [16]. As such, there should be cautious consideration when using the CWIM3 to examine wetlands in and around areas under major influence of anthropogenic land use. It is recommended that, in those cases, to apply a classification using higher-resolution datasets.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

The results of the CWIM3 emphasize the importance of inclusion of climate and ecological information when mapping natural ecosystems at a scale as large as Canada. The Canadian landscape is far from uniform, characterized by mountain ranges, far-reaching plains, forest, and maritime and continental climate areas. The characteristic landscape morphology and climate of distinct ecological areas across Canada (defined by the boundaries of Ecozones) control the formation and expression of wetland distribution, morphology, and vegetation expression. Analysis of classification accuracy results with and without consideration for climate and ecological variation in ecozone reveals the necessity of such datasets for mapping Canada's wetlands. Ecozones can be further broken down into ecoregions [23], areas of even more significant ecological similarity. Integration of ecoregion information in future work may further help to improve wetland accuracy.

Future improvements to the CWIM3 may consider integrating additional satellite data sets such as Hybrid Compact Polarimetry (HCP) data from RADARSAT Constellation Mission (RCM) satellites. Multi-season data has proven to impact smaller-scale wetland classification research accuracy positively and may also be possible. However, there will be some consideration given the difficultly obtaining leaf-off season optical data across the entirety of Canada given issues with cloud cover. This will also increase processing requirements due to a two-fold increase in data inputs. Future work should also integrate additional topographic variables proven to effectively detect wetlands and were not used in the development of the CWIM3, such as the topographic position index [33] and topographic wetness index [34].

Another consideration should be to utilize time-series methodologies such as that performed by [9] to produce low-noise and higher consistency satellite data mosaics. Though perhaps not feasible in the immediate future, an effort to gather wetland test and train data at the sub-class level (wooded bog, wooded fen, shrub swamp, and emergent marsh), etc. may help improve CWIM results. However, most wetland test and train data available to the CWIM are not provided as such, and most are categorized at the level of five classes. Additionally, this will reduce the total number of per-class wetland testing and training data to ingest into the classification methodology.

Future generation of CWIM maps should also focus on improving the accuracy of wetland maps through the application of advanced tools, such as deep learning. Although this may not be possible very soon, as the performance of deep learning tools greatly depend on the availability of large amount of well-distributed training dataset.

VI. CONCLUSIONS

While a problematic endeavor, large-scale wetland classification has become increasingly simplified due to advances in remote sensing satellite data availability, deep learning, and cloud computing. Until recently, Canada has lacked a nationwide data source describing wetland spatial data specifically. Other national data products such as the ACI [1] and the LCC [2] underestimate wetland extent and do not resolve wetlands to the class level. Several generations of the CWIM have been developed to address this problem, improving the results of the previous by integrating new remote sensing data, more significant quantities and quality of training data, and improvements to the RF classification methodology.

Improvements to the CWIM methodology made by the CWIM3 are (1) inclusion of additional remote sensing and auxiliary data including ALOS-2, DEM, nighttime light, climate and precipitation, and alterations to wetland and non-wetland test and train ratios. This has resulted in a ~5 percentage increase in average overall accuracy and reduced wetland class overestimation across all ecozones. This work compares favorably to other research dedicated to determining the wetland extent across Canada [3], [9]. This work demonstrates the importance of multi-source and multi-thematic datasets for wetland classification.

OA's reported by the CWIM3 are higher than that of the CWIM1 and CWIM2, though these values must be interpreted conservatively given the limited distribution of wetland test and training data across certain ecozones, and small number of individual test and train polygons. Increasing wetland test and train data in these areas would certainly increase reliability, though this is not necessarily an attainable goal given funding availability and the isolated nature of many of these ecozones, such as the Taiga Shield. Other issues related to spatial autocorrelation, and the lack of inclusion of topographic variables may also contrite to sources of error within the CWIM3.

Climate change has increased the need for large-scale wetland information, a problem addressed through the development of the CWIM. The CWIM3 represents the highest accuracy Canada-wide wetland classification map, at the level of wetland class, and future research looks to improve these accuracies even more through careful integration of additional multi-source data, and testing and training information.

ACKNOWLEDGMENTS

This research was funded by Natural Resources Canada under Grant (NRCan-3000710039) to M. Mahdianpari. The field data were collected by various organizations, including Ducks Unlimited Canada, Nature Conservancy Canada, Environment and Climate Change Canada (ECCC), Government of Newfoundland and Labrador, and Department of Environment and Conservation, as well as United States National Aeronautics and Space Administration (NASA) and Environmental Protection Agency (EPA). The authors thank these organizations for the generous financial support and for providing such valuable datasets.

REFERENCES

[1] M. Mahdianpari *et al.*, "Meta-Analysis of Wetland Classification Using Remote Sensing: A Systematic Review of a 40-Year Trend in North America," *Remote Sensing*, vol. 12, no. 11, p. 1882, Jun. 2020, doi: 10.3390/rs12111882.

[2] Agriculture and Agri-food Canada, "ISO 19131 Annual Crop Inventory – Data Product Specifications." Government of Canada; Agriculture and Agri-Food Canada; Science and Technology Branch, 2018.

[3] Natural Resources Canada, "2015 Land Cover of Canada." Government of Canada; Natural Resources Canada; Canada Centre for Remote Sensing, 2015.

[4] Fariba Mohammadimanesh, Bahram Salehi, Masoud Mahdianpari, Brian Brisco, and Eric Gill, "Full and Simulated Compact Polarimetry SAR Responses to Canadian Wetlands: Separability Analysis and Classification," *Remote Sensing*, vol. 11, no. 5, p. 516, Mar. 2019, doi: 10.3390/rs11050516.

[5] R. W. Tiner, M. W. Lang, and V. V. Klemas, *Remote sensing of wetlands: applications and advances*. CRC press, 2015.

[6] M. A. Wulder *et al.*, "A National Assessment ofWetland Status and Trends for Canada's Forested Ecosystems Using 33 Years of Earth Observation Satellite Data," *Remote Sensing*, vol. 10, p. 20, 2018.

[7] Environment and Climate Change Canada, "Canadian Environmental Sustainability Indicators Extent of Canada's Wetlands." Government of Canada, 2016.

[8] T. E. Dahl and M. D. Watmough, "Current approaches to wetland status and trends monitoring in prairie Canada and the continental United States of America," *Canadian Journal of Remote Sensing*, vol. 33, no. sup1, pp. S17–S27, 2007.

[9] E. R. DeLancey, J. F. Simms, M. Mahdianpari, B. Brisco, C. Mahoney, and J. Kariyeva, "Comparing Deep Learning and Shallow Learning for Large-Scale Wetland Classification in Alberta, Canada," *Remote Sensing*, vol. 12, no. 1, p. 2, Dec. 2019, doi: 10.3390/rs12010002.

[10] G. Lefebvre *et al.*, "Introducing WIW for detecting the presence of Water In Wetlands with Landsat and Sentinel satellites," *Remote Sensing*, vol. 11, no. 19, p. 2210, 2019.

[11] National Wetlands Working Group, *The Canadian wetland classification system*. Waterloo, Ont.: Wetlands Research Branch, University of Waterloo, 1997. Accessed: Jan. 18, 2021. [Online]. Available: http://www.portofentry.com/Wetlands.pdf

[12] M. Mahdianpari *et al.*, "Smart solutions for smart cities: Urban wetland mapping using very-high resolution satellite imagery and airborne LiDAR data in the City of St. John's, NL, Canada," *Journal of Environmental Management*, p. 111676, 2020.

[13] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, and B. Brisco, "An Assessment of Simulated Compact Polarimetric SAR Data for Wetland Classification Using Random Forest Algorithm," *Canadian Journal of Remote Sensing*, vol. 43, no. 5, pp. 468–484, Sep. 2017, doi: 10.1080/07038992.2017.1381550.

[14] A. Hanson *et al.*, *Wetland ecological functions assessment: an overview of approaches*. Ottawa: Canadian Wildlife Service, 2008.

[15] A. Gallant, "The Challenges of Remote Monitoring of Wetlands," *Remote Sensing*, vol. 7, no. 8, pp. 10938–10950, Aug. 2015, doi: 10.3390/rs70810938.

[16] S. D. Bridgham, J. P. Megonigal, J. K. Keller, N. B. Bliss, and C. Trettin, "The carbon balance of North American wetlands," *Wetlands*, vol. 26, no. 4, pp. 889–916, 2006.

[17] R. A. Fournier, M. Grenier, A. Lavoie, and R. Hélie, "Towards a strategy to implement the Canadian Wetland Inventory using satellite remote sensing," *Canadian Journal of Remote Sensing*, vol. 33, no. sup1, pp. S1–S16, 2007.

[18] A. L. Gallant, *The challenges of remote monitoring of wetlands*. Multidisciplinary Digital Publishing Institute, 2015.

[19] R. C. Gardner and N. C. Davidson, "The ramsar convention," in *Wetlands*, Springer, 2011, pp. 189–203.

[20] D. Mao *et al.*, "National wetland mapping in China: A new product resulting from object-based and hierarchical classification of Landsat 8 OLI images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 164, pp. 11–25, Jun. 2020, doi: 10.1016/j.isprsjprs.2020.03.020.

[21] 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," *Canadian Journal of Remote Sensing*, vol. 46, no. 1, pp. 15–33, Jan. 2020, doi: 10.1080/07038992.2019.1711366.

[22] M. Mahdianpari *et al.*, "The Second Generation Canadian Wetland Inventory Map at 10 Meters Resolution Using Google Earth Engine," *Canadian Journal of Remote Sensing*, vol. 46, no. 3, pp. 360–375, May 2020, doi: 10.1080/07038992.2020.1802584.

[23] Agriculture and Agri-food Canada, "National Ecological Framework for Canada." Government of Canada, Feb. 25, 2013.

[24] M. Mahdianpari et al., "Smart solutions for smart cities: Urban wetland mapping using very-high resolution satellite imagery and airborne LiDAR data in the City of St. John's, NL, Canada," Journal of Environmental Management, p. 111676, 2020.

[25] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18–27, Dec. 2017, doi: 10.1016/j.rse.2017.06.031.

[26] L. Veci, P. Prats-Iraola, R. Scheiber, F. Collard, N. Fomferra, and M. Engdahl, "The sentinel-1 toolbox," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2014, pp. 1–3.

[27] H. K. Zhang and D. P. Roy, "Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification," *Remote Sensing of Environment*, vol. 197, pp. 15–34, 2017.

[28] J. F. Brown *et al.*, "Lessons learned implementing an operational continuous United States national land change monitoring capability: The Land Change Monitoring, Assessment, and Projection (LCMAP) approach," *Remote Sensing of Environment*, vol. 238, p. 111356, 2020.

[29] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[30] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support Vector Machine vs. Random Forest for Remote Sensing Image Classification: A Meta-analysis and

systematic review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020.

[31] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, and M. Motagh, "Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 13–31, Aug. 2017, doi: 10.1016/j.isprsjprs.2017.05.010.

[32] R. Srinet *et al.*, "Mapping plant functional types in Northwest Himalayan foothills of India using random forest algorithm in Google Earth Engine," *International Journal of Remote Sensing*, vol. 41, no. 18, pp. 7296–7309, 2020.

[33] Z. S. Venter and M. A. Sydenham, "Continental-scale land cover mapping at 10 m resolution over Europe (ELC10)," *Remote Sensing*, vol. 13, no. 12, p. 2301, 2021.

[34] Ecosystem Classification Group, *Ecological Regions of the Northwest Territories – Taiga Plains.* Department of Environment and Natural Resources, Government of the Northwest Territories, Yellowknife, NT, Canada., 2009.

[35] D. Phiri, M. Simwanda, S. Salekin, V. R. Nyirenda, Y. Murayama, and M. Ranagalage, "Sentinel-2 data for land cover/use mapping: a review," *Remote Sensing*, vol. 12, no. 14, p. 2291, 2020.

[36] Natural Resources Canada, "National Fire Database Agency Provided Fire Perimeters. 2020." Natural Resources Canada, 2020.

Masoud Mahdianpari (Member, IEEE) received the B.S. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing engineering from the School of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2010 and 2013, respectively, and the Ph.D. degree in electrical engineering from the Department of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, in 2019. In 2019, he was an Ocean Frontier Institute (OFI) Postdoctoral Fellow with Memorial University and C-CORE. He is currently a Remote Sensing Technical Lead with C-CORE and a Cross Appointed Professor with the Faculty of Engineering and Applied Science, Memorial University. His research interests include remote sensing and image analysis, with a special focus on PolSAR image processing, multi-sensor data classification, machine learning, geo big data, and deep learning. Dr. Mahdianpari has authored more than 100 publications, including peer reviewed journal articles, conference papers, books, and book chapters. He was the recipient of the Top Student Award from the University of Tehran for four consecutive academic years from 2010 to 2013, the Research and Development Corporation Ocean Industries Student Research Award, organized by the Newfoundland Industry and Innovation Center, amongst more than 400 submissions, in 2016, the T. David Collett Best Industry Paper Award organized by IEEE, in 2016, The Com-Adv Devices, Inc., Scholarship for Innovation, Creativity, and Entrepreneurship from the Memorial University of Newfoundland, in 2017, the Artificial Intelligence for Earth Grant organized by Microsoft, in 2018, and the Graduate Academic Excellence Award organized by Memorial University, in 2019. He is currently an Editorial Team Member for the journal Remote Sensing and the Canadian Journal of Remote Sensing.

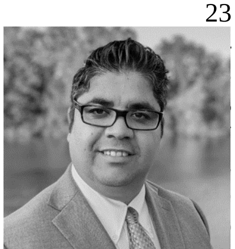
Brian Brisco is an internationally recognized authority on Synthetic Aperture Radar (SAR) and its application to a wide range of environmental monitoring applications. Brian received a Ph. D. degree in remote sensing/physical geography from the University of Kansas in 1985. He also earned a M.Sc. degree in soil science and a B.Sc. degree in ecology, both from the University of Guelph. Brian has been involved in remote sensing since 1975 and participated in the SURSAT project from 1977-1980 before spending 4 years at the Remote Sensing Laboratory at the University of Kansas under the supervision of Dr. F.T. Ulaby, widely recognized as one of the world's leading authorities on radar. He worked for Intera from 1989 until 1997 as a research associate after completion of an NSERC post-doctoral fellowship served at the Canada Centre for Remote Sensing. Since 1997 he has worked for Noetix Research Inc. where he is the Director of Research and Applications Development. His research activities focus on using remote sensing, particularly Synthetic Aperture Radar (SAR), for mapping and managing renewable resources. His work has included experience with interferometry, polarimetry, and radar backscatter modeling including software development and operational implementation. Brian has authored or co-authored more than 200 publications including over 50 peer reviewed journal publications and is the author of two chapters in the Manual of Remote Sensing volume on radar applications published by ASPRS. He provides peer review services to all the major remote sensing journals and participates as an external examiner for many graduate students at various universities in Canada and abroad. Brian has been consulted or contracted by government and non-government organizations on a wide range of SAR applications and system development including NRCan, CSA, DND, AAFC, EC, NASA, ESA, NASDA, NOAA, USDA, CAS, etc. Brian has extensive contacts in the SAR community worldwide and has worked in China, Vietnam, Malaysia, Thailand, Indonesia, South Africa, Argentina, Uruguay, Chile, Brazil, Columbia, and Costa Rica. Dr. Brisco is a Past-President of the Canadian Remote Sensing Society (CRSS) and the Canadian Aeronautics and Space Institute (CASI). Dr. Brisco's skills contribute to ESS Programs on Remote Sensing Science, Earth Science for Environment and Health, Enhancing Resilience in a Changing Climate and Northern Resources and Development.

Jean Granger is a GIS specialist with C-CORE's Earth Observation Team. She received a B.Sc. in Biology in 2012, a diploma in Geographic Information Sciences (GIS) in 2015, and an M.Sc. in Geography in 2019 from Memorial University of Newfoundland (MUN). Her recent work includes contributions to the development of the Canadian Wetland

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Inventory, and wetland inventory maps for the City of St. John's and Miawpukek First Nation.

Fariba Mohammadimanesh received the B.Sc. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing engineering from the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2010 and 2014, respectively, and the Ph.D. degree in electrical engineering from the Department of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, in 2019. She is currently a Remote Sensing Scientist with C-CORE's Earth Observation Team, St. John's, NL, Canada. Her research interests include image processing, machine learning, segmentation, and classification of satellite remote sensing data, including SAR, PolSAR, and optical, in environmental studies, with a special interest in wetland mapping and monitoring, as well as geo-hazard mapping using the interferometric SAR technique (InSAR). Dr. Mohammadimanesh was the recipient of several awards, including the EmeraGraduate Scholarship (2016–2019),CIG-NLAward (2018), and the IEEE NECEC Best Industry Paper Award (2016).



Bahram Salehi is an Assistant Professor of Remote Sensing Engineering at the State University of New York College of Environmental Science and Forestry (SUNY-ESF). He has over 20 years of industrial and academic experiences in applications and algorithm development in Remote Sensing and its environmental applications. His research has resulted in over 50 journal publications and book chapters. At SUNY ESF, his research focus areas are remote sensing of wetlands, forest biomass, and water quality. Dr Salehi has received the 2019 Early Career Medal for his achievements and services in remote sensing from the Canadian Remote Sensing Society. He is a senior member of IEEE, a professional engineer (P.Eng.), and a licensed project management professional (PMP). He is a member of ASPRS, AGU, and CRSS.

Saeid Homayouni (Senior Member, IEEE) received the B.Sc. degree in surveying and geomatics engineering from the University of Isfahan, Isfahan, Iran, in 1996, the M.Sc. degree in remote sensing and geographic information systems from Tarbiat Modares University,Tehran, Iran, in 1999, and the Ph.D. degree in signal and image from Télécom Paris Tech, Paris, France, in 2005. From 2006 to 2007, he was a Postdoctoral Fellow with the Signal and Image Laboratory, University of Bordeaux Agro-Science, Bordeaux, France. From 2008 to 2011, he was an Assistant Professor with the Department of Surveying and Geomatics, College of Engineering, University of Tehran, Tehran, Iran. From 2011 to 2013, through the Natural Sciences and Engineering Research Council of Canada (NSERC) Visitor Fellowship Program, he worked with the Earth Observation Group of the Agriculture and Agri-Food Canada (AAFC), Ottawa Center of Research and Development, Ottawa, ON, Canada. In 2013, he joined the Department of Geography, Environment, and Geomatics, University of Ottawa, Ottawa, ON, Canada, as a Replacing Assistant Professor of Remote Sensing and Geographic Information Systems. Since April 2019, he has been an Associate Professor of environmental remote sensing and geomatics at the Centre Eau Terre Environnement, Institut National de la Recherche Scientifique, Quebec, QC, Canada. He is currently leading a research Group on Earth Observation Analytics by Artificial Intelligence (GEO-AI) with interests in optical and radar earth observations analytics for urban and agro-environmental applications.

Dr. Laura Bourgeau-Chavez is a senior research scientist at Michigan Tech Research Institute and an Adjunct Assistant Professor in the College of Forest Resources and Environmental Science (SFRES) of Michigan Technological University. She has a BS and MS from University of Michigan, School of Natural Resources and Environment as well as a PhD from University of New Brunswick, School of Forestry and Environmental Management. Her main interests are in using passive and active microwave radar imaging for extracting moisture, inundation and biophysical information from wetland and upland landscapes. Her research includes targeted field campaigns to calibrate and validate remotely-sensed data and to better understand ecosystem function and disturbance effects. Her research has been focused on understanding the effects of land use, wildfire disturbance and/or climate change on ecosystems in North American boreal-taiga regions, tropical lowland and mountain peatlands and the coastal Great Lakes. Her work has led to integration of microwave, lidar and optical sensing for mapping and monitoring wetlands, forest soil moisture, forest biomass, invasive species and wildfire effects.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2021.3105645, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

VII. APPENDIX

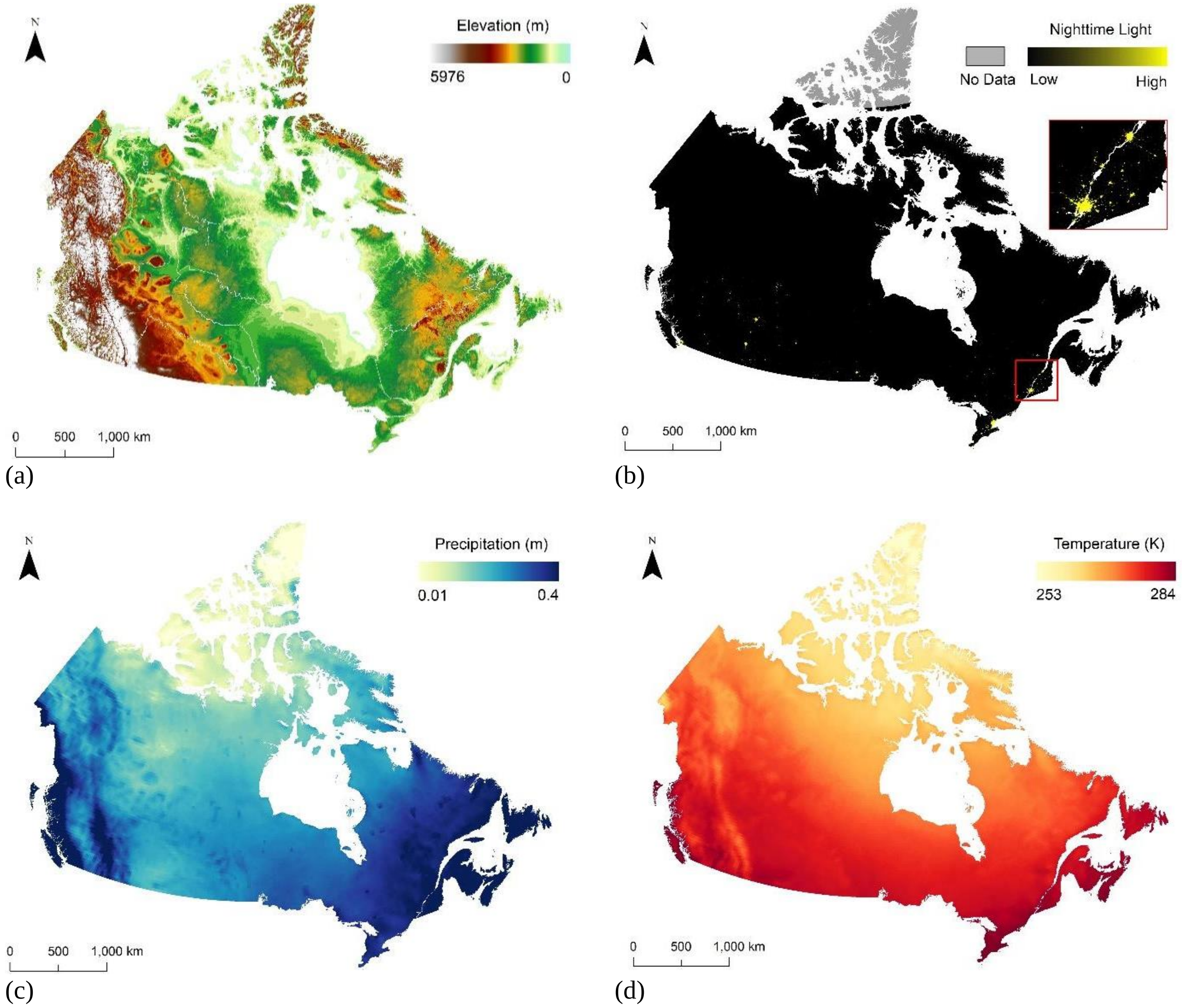


Fig. A. Auxiliary datasets used in the CWIM3 including a Canada-wide 10m digital elevation model (a), nighttime light data (b), precipitation (c), and temperature (d).