

TOWARD A MULTI-SOURCE REMOTE SENSING WETLAND INVENTORY of the USA: PRELIMINARY RESULTS on WETLAND INVENTORY OF MINNESOTA

Sarina Adeli^{1*}, Bahram Salehi¹, Masoud Mahidanpari², Lindi J. Quackenbush¹

(1) Department of Environmental Resources Engineering, College of Environmental Science and Forestry, State University of New York, Syracuse, New York, USA

(2) C-CORE and Department of Electrical Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

KEY WORDS: Geo-Big data, Classification, Machine Learning, Random Forest, Minnesota, Google Earth Engine.

Abstract

Wetlands are highly productive ecosystems that offer unique services on regional and global scales including nutrient assimilation, carbon reduction, geochemical cycling, and water storage. In recent years, however, they are being lost or exploited as croplands due to natural or man-made stressors (1.4 percent in 5 years within the USA). This decline in the extent of wetlands began legislative activity at a national scale that mandate the regulate use of wetlands. As such, the need for cost-effective, robust, and semi-automated techniques for wetland preservation is ever-increasing in the current era. In this study, we developed a workflow for wetland inventorying on a state-wide scale using optimal incorporation of dual-polarimetry Sentinel-1, multi-spectral Sentinel-2 and dual polarimetry ALOS-PALSAR with the Random Forest (RF) classifier in Google Earth Engine (GEE). A total of 45 features from a stack of multi-season/multi-year SAR and Optical imagery (included more than 5000 imagery) was extracted over Minnesota state, USA. We followed the Cowardin classification scheme for clustering the field data. The classification was performed in two levels in 5 different ecozones that cover the Minnesota state. Depending on the availability field data for each ecozone overall accuracies changed from 77% to 85%. The variable importance analysis suggests that Sentinel-2 spectral features are dominant in terms of their capability for wetland delineation. Sentinel-1 backscattering coefficient was also superior among other SAR features. Ultimately, the results of this study shall illustrate the applicability of free of charge earth observation data coupled with the advanced machine learning techniques that are available in GEE for better restoration and management of wetlands.

1. INTRODUCTION

Wetlands offer several significant services on either global and regional scales including carbon sequestration, water purification and weather regulation. However, the rate of wetlands loss is rapidly increasing due to man-made and natural stressors. Dahl et al.(2009) reported that the rate of declining wetlands from 2004 to 2009 is 1.4 percent in the USA (34,050 ha) [1]. To this end, the consistent monitoring of these ecologically important land is essential for their preservation and management. As such, an accurate estimate of wetland's location, extent and their changes through the time is essential. The availability of fine- resolution Earth Observation (EO) on a sub-weekly basis coupled large-scale computing capability of Google Earth Engine (GEE) as well as advanced machine learning tools can facilitate wetland mapping [1], [2]. In particular, the optimal incorporation of optical and Synthetic Aperture Radar (SAR) EO data can enhance the wetland delineations since they are sensitive/responsive to different wetland characteristics [3], [4]. The ultimate goal of this study is to propose a fast, cost-effective, robust, and semi-automated technique for wetland classification using cloud computing platforms and multi-source EO data with Random Forest (RF) classifier.

2. METHOD

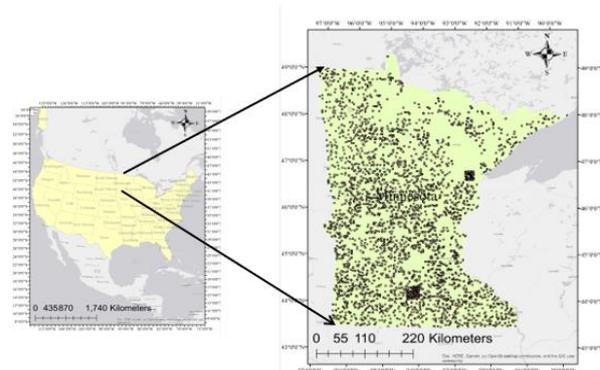


Figure 1. The geographic location of case study and the distribution of reference data

This study focuses on Minnesota state. Minnesota is the 12th largest state in the USA with an area of 225181.336 square kilometres. Minnesota is divided into 5 different ecozones. These zones presumably represent similar weather, vegetation and, ecological patterns. The distribution of reference data for the classification is shown in (Figure 1.). We followed the Cowardin classification scheme that also NWI is adopting for generating wetlands inventory maps [4].

*Corresponding Author

Based on the updated Cowardin system the considered wetland classes for this study are emergent, scrub/shrub, and forested wetlands. Figure 2. represents the number of acquired data for each ecozones. An optimal fusion of Sentinel-1, Sentinel-2 and, ALOS-PALSAR imageries was created for each ecozones. We also added slope and aspect features that were derived from SRTM 30-Meter DEM. Given the cloudy climate of Minnesota, we applied cloud masks on optical Imageries. Further, a 3 years summer composite from June to August was created to assure wall to wall coverage of all the state. Speckle was also removed from SAR imageries using a circular Kernel-based filter with the radius of 25. Lastly, all the available data was resampled to 10 meters using the bilinear resampling technique.

Once the data was pre-processed, the desired features from more than 6000 satellite imagery were extracted. We calculated vegetation, water and, soil indices for optical imageries. The calculated SAR features include ratio, power, span or total power and backscattering coefficients. Further by stacking the extracted features into one single vector with dimension of 28, we applied the RF classifier in two levels. After performing several tests, the optimal number of trees for the RF classifier was set to 400. The number of splitting nodes was set to square root of the number of features. The reference data was separated to 70% for train and 30% for test using random resampling. The accuracy assessment and variable importance analysis using independent test data were implemented once the RF classifier was applied.

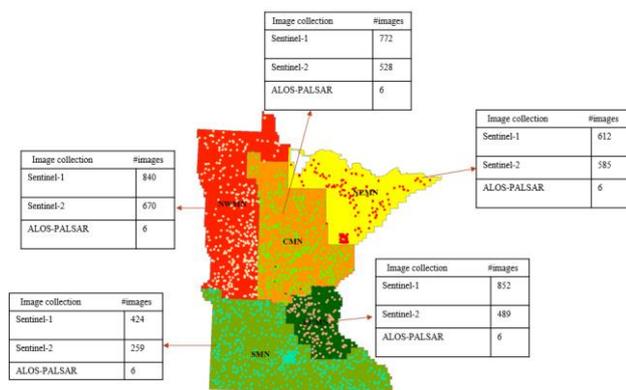


Figure 2. Number of acquired imageries and distribution of test data for each ecozone.

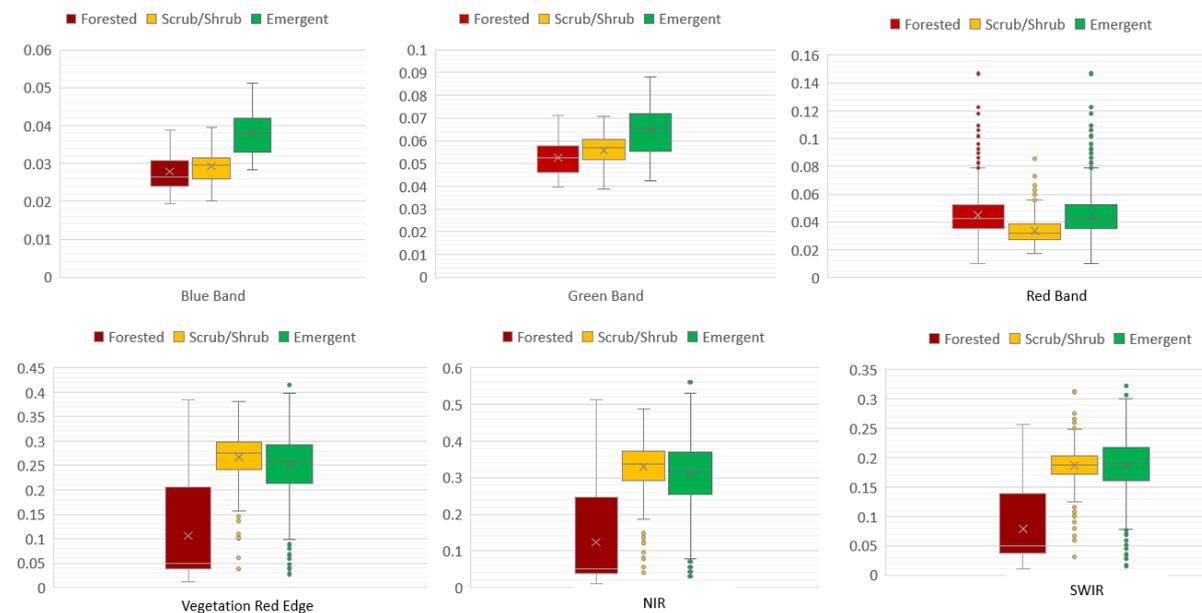


Figure 3. Box-and-whisker plot of the multi-year/ multi-season composite illustrating the distribution of training data reflectance.

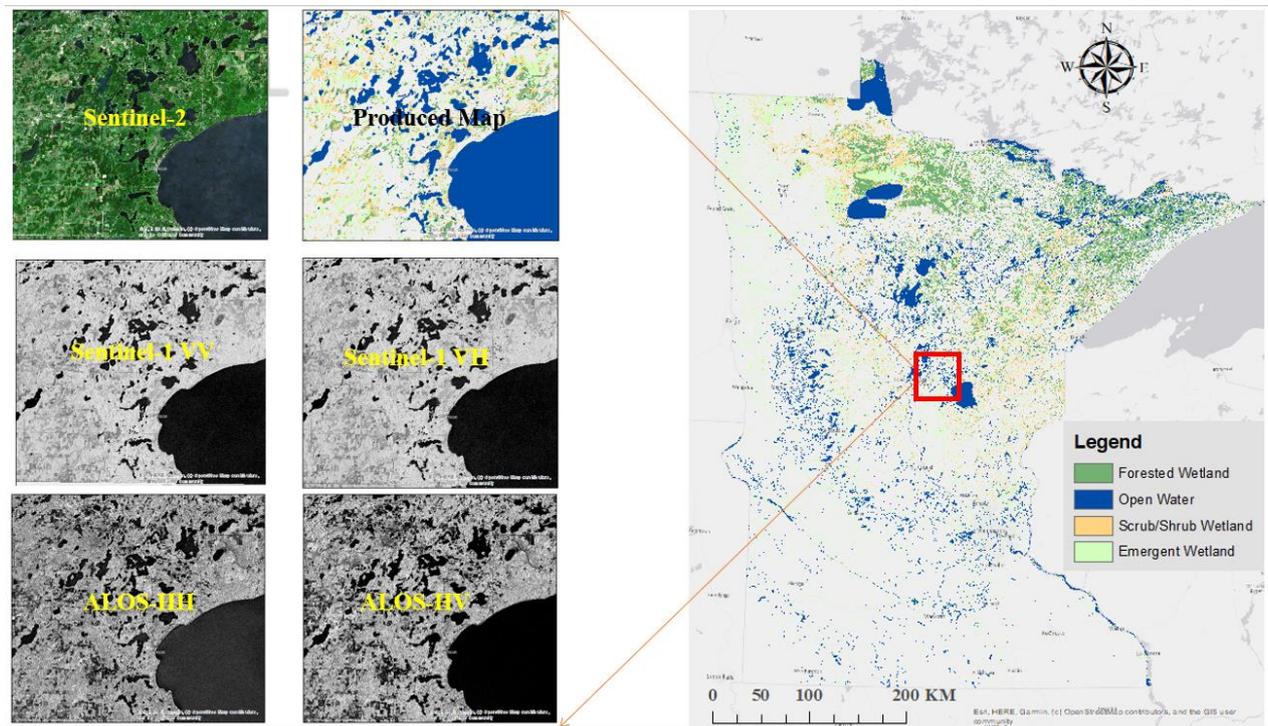


Figure 4. Right: the classified map for the entire state; Left: the zoom-in scenes of the red rectangle.

The box-and whisker plots of Red, Green, Blue, Vegetation Red Edge, NIR and SWIR bands of Sentinel-2 for three class of wetlands are illustrated in Figure 3. The initial reflectance separability analysis suggests that emergent and scrub/shrub

wetland's separation is maximized in blue band. Moreover, the separation of forested wetland from the other two classes of wetland is augmented at NIR, SWIR and Vegetation Red Edge bands.

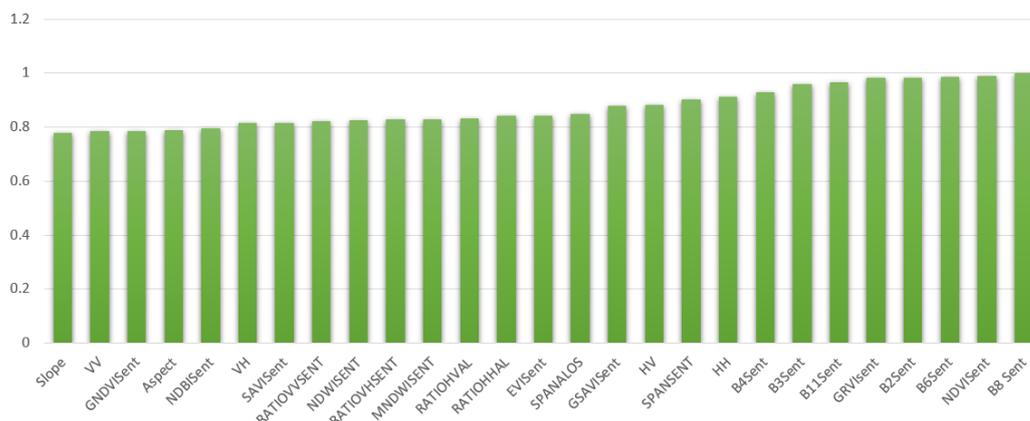


Figure 5. Variable importance analysis of RF classifier.

3. RESULTS AND DISCUSSION

Figure 4 shows the classification results over the Minnesota state on the left. The zoom-in version of the produced map and used satellite imagery is shown in the left. Notably, the upland class is excluded from the final classified maps since we were keen to examine the wetland's distribution.

Following, post analysis accuracy assessment using independent test data reveals that the level one classification wetland's producer's accuracy varies between 80.75% and 88.23% for Central and Northwest ecozones, respectively. The wetland's user's accuracy varies from 78.66% to 87.37% for Northeast and Northwest, respectively. The overall wetland-upland accuracy changes from 83.66% to 90.90% for Northeast and Northwest ecozones. The level

two classification overall accuracies that correspond only to wetlands classes vary from 63.12% to 71.25% for Central and South, respectively. These changes in the accuracies are potentially due to the insufficient train data for some eozones. Further, variable importance analysis of 28 imported feature reveals the superiority of Sentinel-2 spectral bands and derived indices (Figure 5). The first SAR feature that emerged was the backscattering coefficient of horizontally transmitted and horizontally received Sentinel-1 data. This is consistent with other studies as this feature is explicitly powerful in delineating herbaceous wetlands classes [3]. Overall, the results illustrated in this study shall provide some initial insight on the potential ability of EO geo-big data with 10-meter resolution using GEE platform as well as advance machine-learning technique for wetland mapping on a state-wide scale. The produced maps can be updated in sub-weekly basis that is an asset for dynamic wetland environment.

4. ACKNOWLEDGEMENTS

Field data, used in this study, was generously provided by Steven Kloiber from Minnesota Department of Natural Resources.

4. REFERENCES

- [1] T. E. Dahl et al., "Status and Trends of Wetlands in the Conterminous United States 2004 to 2009," p. 112, 2009.
- [2] 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, doi: 10.1080/07038992.2019.1711366.
- [3] S. Adeli *et al.*, "Wetland Monitoring Using SAR Data: A Meta-Analysis and Comprehensive Review," *Remote Sens.*, vol. 12, no. 14, Art. no. 14, Jan. 2020, doi: 10.3390/rs12142190.
- [4] L. M. Cowardin, V. Carter, F. C. Golet, and E. T. LaRoe, "Classification of wetlands and deepwater habitats of the United States," U.S. Department of the Interior, U.S. Fish and Wildlife Service, Technical Report, 1979. Accessed: Oct. 15, 2019. [Online]. Available: <https://tamug-ir.tdl.org/handle/1969.3/20>