

# MAPPING PLANT COMMUNITIES IN THE INTERTIDAL ZONES USING SENTINEL-2 AND SENTINEL-1 DATA

Tiejun Wang<sup>1</sup>, Yansha Luo<sup>1,2</sup>, Yiwen Sun<sup>1\*</sup>, Xinhui Liu<sup>2</sup>

<sup>1</sup> University of Twente, 7500 AE Enschede, The Netherland

<sup>2</sup> Beijing Normal University, Beijing 100875, China

t.wang@utwente.nl, yansha\_luo@163.com, y.sun@utwente.nl, xhliu@bnu.edu.cn

## ABSTRACT

Intertidal habitats are important not only for the ecosystem services they provide but also because they represent a key ecotone between land and ocean. In this paper we explore the synergistic use of time series Sentinel-1 and Sentinel-2 data for mapping intertidal plant communities, using a Random Forest algorithm. We compare the performances of ten models with different input data. Results showed that the use of multi-seasonal Sentinel-2 images could significantly improve the mapping accuracy compared to the use of any single-season image. There was no statistically significant difference in mapping accuracies between the use of multi-seasonal and time series Sentinel-2 images. However, when combining time series Sentinel-1 data, time series Sentinel-2 data and NDVI statistic metrics, the highest mapping accuracy was achieved with an overall mapping accuracy of 77.7% and the Kappa coefficient of 0.75.

**Index Terms**— Intertidal plant communities, Sentinel-1, Sentinel-2, Single-season, Multi-season, Random forest

## 1. INTRODUCTION

The intertidal zones represent the areas which are above water at low tide and under water at high tide. This ecotone is commonly regarded as one of the most valuable and active wetland ecosystems because of its rich biodiversity, economic value, and ecological functions [1]. Intertidal zones provide important habitats for a wide variety of marine and terrestrial life [2]. Due to daily tidal cycles, intertidal zones are characterized as highly dynamic environments, which cause variation of moisture, temperature, nutrients and soil salinity [3]. These extreme environments influence marine and terrestrial biodiversity through food chains and interactions. However, wetland ecosystems, including intertidal ecosystem, are globally

threatened and degraded due to increasing anthropogenic pressure, climate change and consequent sea-level rise.

Plant communities are the key components of intertidal zones. They can be used to indicate the health of ecosystems, providing early signs of ecological degradation [4]. Plant communities in wetlands, having a capacity to remove toxic substance and heavy metals, are therefore regarded as an indicator of water pollution. Moreover, intertidal plant communities provide key habitats for migratory birds and prevent coastlines from soil erosion, playing a crucial role in shore protection and sediment retention [1]. Mapping and monitoring of the plant communities in intertidal zones have become increasingly important to provide basic data for sustainable environmental management.

Field accessibility hinders ground surveys for mapping the distribution of intertidal plant communities over large areas. Satellite-based remote sensing provides an ideal platform for land surface monitoring in term of both cost and coverage, which has been commonly used for mapping vegetation and its associated biophysical and ecological parameters [5]. However, remotely distinguishing and mapping the intertidal plant communities is a challenging task due to the complexity and highly dynamic characters of this landscape.

The recent deployment of ESA's Sentinel operational satellites has established a new paradigm for remote sensing applications. Sentinel-2 optical imagery has great potential for mapping intertidal plant communities thanks to its improved spatial and temporal resolution as well as three red edge bands that are sensitive to chlorophyll content. Optical remotely sensed time series data had been widely applied in vegetation classification [6]. Multi-temporal images may capture the seasonality or phenological variation in vegetation types and therefore improve the accuracy of distinguishing different vegetation types. However, cases were also found that single-season image can produce higher mapping accuracy in some landscapes [7]. In addition to optical imagery, the synthetic aperture radar (SAR) data has also been increasingly explored for mapping and monitoring

---

\*To whom correspondence should be addressed.

wetland vegetation [8]. SAR microwaves can penetrate through clouds and vegetation (depending on frequency) and can operate in day or night conditions. In this context, Sentinel-1 radar images have made it possible to classify different land cover types with a high spatial and temporal resolution. The study aimed to explore the synergetic use of time series Sentinel-1 SAR and Sentinel-2 optical data for mapping plant communities in the intertidal zones of the Yellow River Delta, China. Specifically, we set out to verify 1) if multi-seasonal Sentinel-2 images can significantly improve the classification accuracy of the intertidal plant communities compared to a single-season Sentinel-2 image; 2) can the combination of time series Sentinel-1 and Sentinel-2 data produce the highest mapping accuracy for the intertidal plant communities?

## 2. MATERIALS AND METHODS

### 2.1. Study area

The study area (37°32'N - 37°54'N, 118°39'E - 119°21'E) is located in the intertidal zones of the Yellow River estuary, which is part of the Yellow River Delta National Reserve situated in the northeast of Dongying City, Shandong Province, China. There are four dominant plant species: smooth cordgrass (*Spartina alterniflora*), seepweed (*Suaeda salsa*), common reed (*Phragmites australis*), and five-stamen tamarisk (*Tamarix chinensis*). Based on literature and local expert knowledge, seven classes of plant communities and two classes of non-vegetation types in this area were identified (see Table 1).

Table 1. Seven classes of plant communities and two types of non-vegetated areas identified in the study area along with the associated sample size used for training and testing.

Land cover classes	Training samples	Testing samples
Seepweed	23	23
Seepweed + reed	17	17
Seepweed + tamarisk	20	20
Seepweed + reed + tamarisk	25	25
Reed	17	17
Reed + tamarisk	16	16
Cordgrass	19	19
Mudflats	25	25
Open water	23	23
Total	185	185

### 2.2. Data collection and pre-processing

Twenty-four Sentinel-1 SAR images (dual VV+VH polarisation) and twelve Sentinel-2 optical images (bands 2-7, 8a, 11 and 12) were downloaded from the Copernicus Open Access Hub (<https://scihub.copernicus.eu/dhus/>). The

Sentinel-1 images were acquired on 19 July, 29 September, 23 October, 28 November and 22 December of 2016, and 27 January, 20 February, 28 March, 21 April, 27 May, 8 June and 26 July of 2017, respectively. The Sentinel-2 images were acquired on 16 June, 24 September, 14 October and 3 November of 2016, and 12 January, 11 February, 2 April, 22 April, 12 May, 1 June, 21 June and 11 July of 2017, respectively.

Image pre-processing was carried out using the Sentinel Toolbox in the Sentinel Application Platform (SNAP), including radiometric calibration, speckle filtering, terrain correction for Sentinel-1 data and atmospheric correction for Sentinel-2 data. All images were resampled to 20 m resolution. The red-edge region (700-740nm) has been introduced to Sentinel-2 data to increase sensitivity to small changes in canopy foliage content, gap fraction and senescence, which might be useful for the mapping of intertidal plant communities. Therefore, a red edge band based normalised difference vegetation index (NDVI) [ $NDVI_{705} = (B06 - B05) / (B06 + B05)$ ] was also used in this study. As a result, four statistical metrics (i.e., maximum, minimum, mean, and standard deviation) were calculated based on the time series NDVI dataset. Ground observation data for image classification and verification were collected in August 2017, using a stratified random sampling method. For each class, at least 30 samples were collected (see Table 1) with a plot size of 30×30 m. In total, 370 samples were collected and used in this study.

### 2.3. Random Forest classification algorithm

Random Forest is a widely used machine learning algorithm and has become popular within the remote sensing community on account of its competitive performance in image classification [9]. In this study, the number of classification trees desired was set equal to 1000 and the number of prediction variables used in each node to grow the tree was set equal to the square root of the total number of input features. Ten different combinations of input remotely sensed data were explored (Table 2). We made ten random partitions of the samples by randomly dividing into two equal groups for training and testing the classification. The models were then run on each partition to reduce the potential bias introduced by the samples. All the analyses were conducted with the R package 'randomForest'.

### 2.4. Accuracy assessment and model comparison

The accuracy of the classification maps derived from different combinations of input data were assessed using a confusion matrix. The overall accuracy and kappa coefficient were calculated. The statistical significance of differences in accuracies achieved by classifications with different combinations of input data was evaluated using a pairwise two-sample t-tests.

Table 2. Ten models with different input data.

Models	Input data						
	Sentinel-2				Sentinel-1		
	Spring	Summer	Autumn	Winter	Time series images	NDVI statistic metric	Time series VV&VH
1	✓						
2		✓					
3			✓				
4				✓			
5	✓	✓	✓	✓			
6					✓		
7						✓	
8					✓	✓	
9							✓
10					✓	✓	✓

### 3. RESULTS

#### 3.1. Mapping of intertidal plant communities

The classification maps of intertidal plant community varied among ten models. The map with the highest overall accuracy and kappa coefficient is shown in Figure 1.

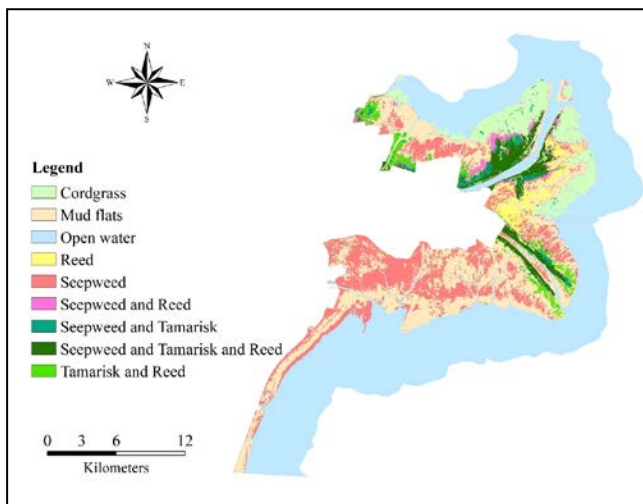


Figure 1. The classification map with the highest accuracy (overall accuracy=77.7%; kappa=0.75) derived from the Model 10: the combination of the time series Sentinel-1 data, time series Sentinel-1 data and NDVI statistical metrics.

The areas covered by cordgrass are mainly found in river estuary and along the north coastline of the delta. From the river bank to the sea, plant succession caused by soil conditions could be recognized. The plant communities showed a sequential structure, from seepweed + tamarisk + reed to seepweed + tamarisk, and to seepweed + reed, and followed by seepweed as well as cordgrass. This distribution

pattern appears to be associated with the gradient of soil salinity and the salt tolerance of different plant species. The plant community of the seepweed + tamarisk + reed was mostly distributed along the freshwater, within the high tide zone. In the south part of intertidal zones, the plant community was less diverse, covering by a large extent of mudflats combined with seepweed.

#### 3.2. Accuracy assessment and model comparison

The mapping accuracies (overall accuracy and kappa coefficient) derived from the ten classification models are shown in Figures 2 and 3.

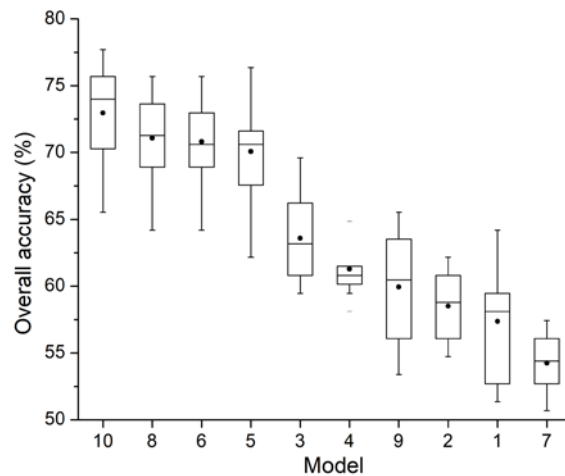


Figure 2. Boxplots show the overall accuracies of the ten models with ten repetitions.

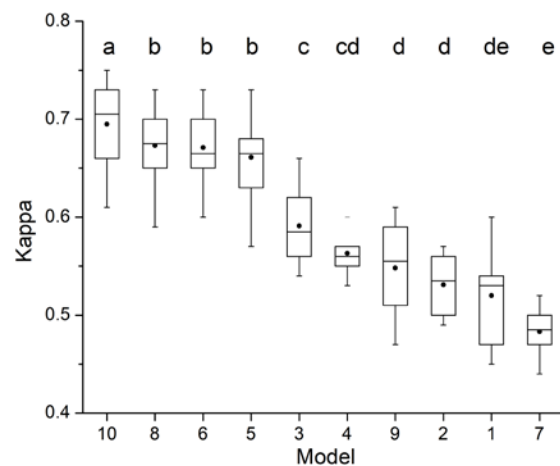


Figure 3. Boxplots show the kappa coefficients of the ten models with ten repetitions. Different letters on boxes denote significant difference at  $p < 0.05$  level.

The mapping accuracies of the single-season images captured in spring (overall accuracy=57.36%; kappa

coefficient=0.52), as well as in summer (overall accuracy=58.51%; kappa coefficient=0.53) were significantly lower than the results derived from autumn image (overall accuracy=63.58%; kappa coefficient=0.59) and winter image (overall accuracy=61.28%; kappa coefficient=0.56). The season of autumn performed best among all seasons due to the distinctive phenological characteristics of various vegetation species.

The mapping accuracies derived from the multi-seasonal Sentinel-2 images (overall accuracy=70.07%; kappa coefficient=0.66) were significantly higher (t-test,  $p<0.05$ ) than the mapping accuracies derived from any of the four single-season Sentinel-2 images. So the use of multi-seasonal Sentinel-2 images can significantly improve the mapping accuracy of the intertidal plant communities compared to a single-season image.

The mapping accuracies derived from the time series Sentinel-2 images (overall accuracy=70.81%; kappa coefficient=0.67) were slightly higher than the one derived from the multi-seasonal Sentinel-2 images (overall accuracy=70.07%; kappa coefficient=0.66). However, there was no statistically significant difference between them (t-test,  $p>0.05$ ). A plausible explanation seemed to be that the multi-seasonal Sentinel-2 images contained most variable importance for mapping vegetation in such dynamic wetlands and the extra images did not contribute so much to the classification of time series images.

The NDVI statistical metrics derived from the time series Sentinel-2 data have no sufficient capability for mapping the intertidal plant communities by itself, resulting in the worst performance with an overall accuracy of 54.26% and kappa coefficient of 0.48, which was significantly lower (t-test,  $p<0.05$ ) than the mapping accuracies derived from any of the other nine models. Also, the mapping accuracies derived from the combination of the time series Sentinel-2 data and NDVI statistic metrics (overall accuracy=70.81%; kappa coefficient=0.67) were not statistically higher (t-test,  $p>0.05$ ) than the one derived from the time series Sentinel-2 data alone (overall accuracy=70.81%; kappa coefficient=0.67). The reason may lie in the confusion of mixed plant communities within one pixel. Moreover, adding the NDVI statistical metrics to the time series Sentinel-2 images achieved a mean overall accuracy of 70.81% with a mean kappa coefficient of 0.67, which did not significantly improve the mapping accuracy than the model without the use of the NDVI statistical metrics.

The mapping accuracies derived from the combination of the time series Sentinel-1 SAR and Sentinel-2 optical images (overall accuracy=72.97%; kappa coefficient=0.70) were statistically higher (t-test,  $p<0.05$ ) than the mapping accuracies derived from either the time series Sentinel-2

optical images (overall accuracy=70.81%; kappa coefficient=0.67) or the time series Sentinel-1 SAR images (overall accuracy=59.93%; kappa coefficient=0.55).

#### 4. CONCLUSION

This paper explored the synergistic use of the time series Sentinel-1 SAR and Sentinel-2 optical images for mapping intertidal plant communities. The results confirm that the use of multi-seasonal Sentinel-2 images can significantly improve the mapping accuracy of the intertidal plant communities compared to a single-season Sentinel-2 image. There is no statistically significant difference in intertidal plant communities mapping accuracies between the use of time series Sentinel-2 data with and without the use of NDVI statistic metrics. The combination of the time series Sentinel-1 and Sentinel-2 data can significantly improve the mapping accuracy of the intertidal plant communities compared to the use of any of the single sensor time series data. However, when combining Sentinel-1 time series, Sentinel-2 time series and NDVI statistic metrics, the highest mapping accuracy was achieved with an overall mapping accuracy of 77.7% and the Kappa coefficient of 0.75.

#### 5. REFERENCE

- [1] R. Costanza et al., "The value of the world's ecosystem services and natural capital," *Ecological Economics*, vol. 25, no. 387, pp. 3–15, 1998.
- [2] S.M. Rog and C.N. Cook, "Strengthening governance for intertidal ecosystems require a consistent definition of boundaries between land and sea," *Journal of Environmental Management*, vol. 197, pp. 694–705, 2017.
- [3] B.A. Menge and G.M. Branch, "Rocky intertidal communities," in M.D. Bertness et al. (Ed.) *Marine community ecology*, Sinauer Associates, Sunderland, pp. 221–251, 2001.
- [4] W.C. Dennison et al., "Assessing water quality with submersed aquatic vegetation," *Bioscience*, pp. 86–94, 1993.
- [5] Y. Xie, Z. Sha and M. Yu, "Remote sensing imagery in vegetation mapping: a review," *Journal of Plant Ecology*, vol. 1, no. 1, pp. 9–23, 2008.
- [6] C. Gómez, J.C. White, and M.A. Wulder, "Optical remotely sensed time series data for land cover classification: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 55–72, 2016.
- [7] S.K. Landley, H.M. Cheshire and K.S. Humes, "A comparison of single date and multitemporal satellite image classifications in a semi-arid grassland," *Journal of Arid Environments*, vol. 49, no. 2, pp. 401–411, 2001.
- [8] L. White et al., "A collection of SAR methodologies for monitoring wetlands," *Remote Sensing*, vol. 7, no. 6, pp. 7615–7645, 2015.
- [9] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.