

RANDOM FOREST CLASSIFICATION OF SEDIMENTS ON EXPOSED INTERTIDAL FLATS USING ALOS-2 QUAD-POLARIMETRIC SAR DATA

W. Wang^{a,b,*}, X. Yang^a, G. Liu^a, H. Zhou^{a,b}, W. Ma^a, Y. Yu^a, Z. Li^a

^a RADI, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China -
(wangws, yangxf, liugh, zhouhh, mawt, yuyang, lizw)@radi.ac.cn

^b UCAS, University of Chinese Academy of Sciences, Beijing 100101, China

Commission VIII, WG VIII/9

KEY WORDS: Coastal Zones Surveillance, SAR, Polarimetric Decomposition, Optical Channels, Random Forest

ABSTRACT:

Coastal zones are one of the world's most densely populated areas and it is necessary to propose an accurate, cost effective, frequent, and synoptic method of monitoring these complex ecosystems. However, misclassification of sediments on exposed intertidal flats restricts the development of coastal zones surveillance. With the advent of SAR (Synthetic Aperture Radar) satellites, polarimetric SAR satellite imagery plays an increasingly important role in monitoring changes in coastal wetland. This research investigated the necessity of combining SAR polarimetric features with optical data, and their contribution in accurately sediment classification. Three experimental groups were set to make assessment of the most appropriate descriptors. (i) Several SAR polarimetric descriptors were extracted from scattering matrix using Cloude-Pottier, Freeman-Durden and Yamaguchi methods; (ii) Optical remote sensing (RS) data with R, G and B channels formed the second feature combinations; (iii) The chosen SAR and optical RS indicators were both added into classifier. Classification was carried out using Random Forest (RF) classifiers and a general result mapping of intertidal flats was generated. Experiments were implemented using ALOS-2 L-band satellite imagery and GF-1 optical multi-spectral data acquired in the same period. The weights of descriptors were evaluated by VI (RF Variable Importance). Results suggested that optical data source has few advantages on sediment classification, and even reduce the effect of SAR indicators. Polarimetric SAR feature sets show great potentials in intertidal flats classification and are promising in classifying mud flats, sand flats, bare farmland and tidal water.

1. INTRODUCTION

The intertidal zone is the coastal area between the daily main high waterline and the daily mean low waterline (Gade et al., 2014), and falls dry once during every tidal cycle. A frequent surveillance of intertidal regions is mandatory due to national and international directives. Sediments on intertidal flats perform vital ecosystem functions that play important roles in regulating global hydrology and climate; modifying chemical elements of water column, sediment, macrophyte and biotic communities; and protecting shorelines from erosion and storms. However, traditional ways of monitoring the health conditions of sediment flats are always difficult tasks because of the repetitive flooding and shallow depths of water. Remote sensors can act as frequent and rapid monitors and being widely used.

Optical remote sensing are already being used for sediment extraction and habitat mapping on exposed intertidal flats (Brockmann and Stelzer, 2008). With the advent of microwave sensors, the use of SAR images is becoming a hot spot because of their all-weather capabilities and independence of daylight (Gade et al., 2015). Former studies have thrown light on the potentials of the features extracted from optical and SAR data source and abilities of discrimination of tidal water, sediment and other biotic community. (Baghdadi et al., 2013; Banks et al., 2014a; Banks et al., 2014b; Brockmann and Stelzer, 2008; Gade et al., 2008; Koch et al., 2012; Ullmann et al., 2014). But there

are rarely studies focus on subdividing sediment types which can promote economic development and environmental protection. Various sediment types including mud flats, sand flats, bare farmland perform different ecosystem function and provide nursery habitat for various aquatic organisms. In addition, the feasibility of these data source for sediment interpretation has not yet been systematic evaluated. It is also essential to assess whether the combinations of optical and SAR images can perform unique advantages on frequent and accurate observation of sediment targets. Furthermore, different classifiers are widely used in different classification scheme such as complex Wishart classifier and Fuzzy C-means classifier. These algorithms strongly depend on parameter setting and are very time-consuming. Random Forest (RF) classifier is used in this paper because of its non-parameter attribution and time-efficient ability (Breiman, 2001). Many researchers have applied successfully in ecological land cover studies (Corcoran et al., 2012; Rodriguez-Galiano et al., 2012; van Beijma et al., 2014). It is especially well suited for classification of multi-source data and geographical data (van Beijma et al., 2014). Based on the above motivations, we want test the behaviour of different feature sets using RF classifier.

In this paper, we propose systematic assessment of features from different RS data source when used in sediment classification on exposed intertidal flats. Study site descriptions and classification schemes based on different feature

* Corresponding author: Wensheng Wang, (wangws@radi.ac.cn) current Ph.D. candidate of RADI, CAS.
Main research direction is Ocean Remote Sensing, SAR Oceanography.

combinations are introduced in the methodology part. Then, result analysis and conclusions are elaborated in the following part. This paper gives essential supports for further study on remote sensing of intertidal regions.

2. METHODOLOGY

This section includes the introduction of study site and data source we use in this paper; the explanation of pre-processing, feature sets extraction and image classification.

2.1 Study Site and Data Source

The study site is chosen in Ganga Delta which locates on the border between south-western China and north-eastern India centered at 27°02'N and 94°00'E. This alluvial plain is generated by the repetitive flooding of Brahmaputra Rivers (see Figure.1 (left)). The target area contains water system, mud flats, sand flats, man-made structures, forests, and bare soil from cropland. All our target objects have wide distribution which correspond well with the objectives in this paper.

As for SAR data source, we use Advanced Land Observation Satellite-2 Phased Array type L-band SAR (ALOS-PALSAR-2) level-1.1 data that was acquired on Nov. 17th in 2015. The pixel spacing in slant azimuth direction is 2.78m while that in slant range direction is 2.86m. Totally four channels including HH, HV, VH, and VV with SLC format were acquired. As shown in Figure.1 (right), HH, HV, VV and their Pauli (R-HH, G-HV, B-VV) colour composite thumb images are displayed. As for optical data source, we use Level 1A multispectral images from GF-1 satellite which is the first satellite of the high resolution earth observation satellite system in China. It was acquired on Nov. 17th in 2015 with spatial resolution 8m in multispectral bands and 2m in panchromatic band. Both the SAR and optical data source are resized to 8m before pre-processing. The validation dataset is generated by artificial visual interpretation results from the fused optical data with 2m spatial resolution

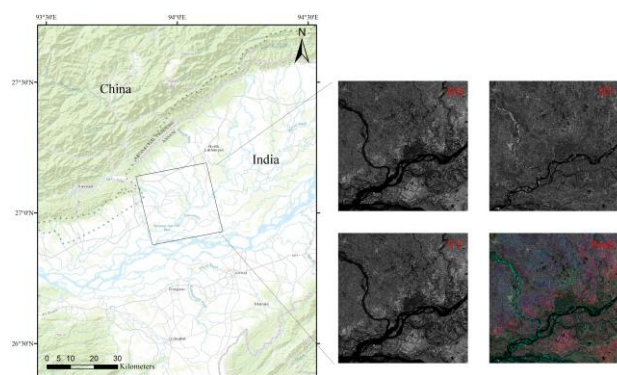


Figure 1. Overview map of study area (left) and HH, HV, VV and Pauli colour composite images from ALOS-2 satellite.

2.2 Data Processing

2.2.1 SAR Data Pre-processing: SAR data source from ALOS-2 is pre-processed by multilooking filter of 2 in azimuth direction and converted from slant to ground range. To reduce the speckle noises caused by interference between different bound to microwave, a refined Lee filter is performed with the function of protecting the details and texture information to the

largest extent. The GF-1 data with Red, Green, and Blue channels are pre-processed by RPC ortho-rectification and masking to guarantee the same geographic range with SAR data. The two data source are then converted to the WGS-84 projection and resized to 8m \times 8m spatial resolution.

2.2.2 Feature Sets Generations: Freeman-Durden (FD), Cloude-Pottier (CP) and Yamaguchi decomposition methods are commonly applied in target coherence decomposition (TCD). FD method performs on the basis of the radar scattering mechanism and decomposes the backscattering observation of natural media surfaces into three physical scattering mechanisms (Freeman and Durden, 1998). The components from FD contains FD_surface, FD_volume, and FD_double-bounce. As for the CP method, it performs on the basis of eigenvalues and corresponding eigenvectors, and divides the feature-space to identify different targets with different scattering mechanism (Cloude and Pottier, 1996). The components from CP contains CP_entropy, CP_anisotropy, and CP_alpha. Yamaguchi method has four components including Y_double bounce, Y_volume, Y_surface, and Y_helix (Koch et al., 2012). These three mainstream TCD methods have been demonstrated high feasibility in wetland classification. The objective of this paper is to assess the contribution of these SAR features and demonstrate whether the simplification of these features has the ability of distinguishing the sediment types on exposed intertidal flats.

Three bands of the GF-1 optical data with Red, Green, and Blue channels are considered using their colours and spectral characteristics. The combination of SAR and optical data is used to evaluate whether optical data source can increase the classification precision of sediment targets, and to see whether single SAR data source can be competent for sediment discrimination. After all, the different imaging time gives errors when it comes to these rapid changing and unstable shaped targets.

2.2.3 Image Classification: Considering the objective of this paper, we need to assess the potentials of SAR descriptors when used for intertidal flats discrimination, and evaluate whether SAR data source can get better classification accuracy of sediments comparing with multi-source dataset. We set three experimental groups (see Table. 1): (i) The SAR features extracted from FD, CP and Yamaguchi methods, totally 10 features with electromagnetic and structure information; (ii) The optical channels including Red, Green, and Blue intensity with spectral information; (iii) The combination of SAR and optical features, totally 13 features are prepared to following classification. The classification is process by RF classifier. RF classifier is based on decision tree classification and performed by fitting a predefined number of classification trees (in this research 500) to a feature set. RF variable importance indexes are calculated to assess the contribution of each feature. The classification results are given by overall accuracy, user accuracy, producer accuracy and RF variable importance indexes.

	GROUP(i)	GROUP(ii)	GROUP(iii)
Freeman-Durden	✓		✓
Cloude-Pottier	✓		✓
Yamaguchi	✓		✓
R/G/B channels		✓	✓

Table 1. Experimental group settings

3. RESULTS

This research investigated the necessity of combining SAR polarimetric features with optical data, and their contribution in accurately sediment classification. Three feature combinations are processed using RF classifier and generated sediment classification maps (See Figure. 2). The results are assessed using confusion matrices, and the accuracies of certain classes are listed in a table (See Table. 2). As the confusion matrices show, the group (i) with SAR indicators performed with the highest accuracy; group (ii) gave the worst overall accuracy; group (iii) performed with mix success. Obviously, group (i) performed best in mud flats, sand flats, and bare croplands classification while group (ii) conversely failed to classify these sediment types. However, group (ii) showed its potential in vegetation extraction with the higher single class accuracy than SAR feature set. Group (iii) contains SAR and optical channels gave the highest accuracy in vegetation classification but decreased the effect of SAR indicators in sediment classification. It indicated that multi-source feature can distinguish targets with significantly different scattering characteristics, but for targets with similar scattering types, optical channels always get in the way.

From the result of RF variable importance indexes, not all SAR indicators in this paper are suitable for sediment subdivision. After calculating the importance indexes, we can see that Yamaguchi components are not that suitable for sediments with similar radar backscattering signatures. Y_surface, Y_double-bounce, and Y_helix ranked behind the sorted list. In particular, the fourth components of Yamaguchi (helix scatters) takes into account co- and cross-pol correlations that are relevant for man-made targets but less relevant in a natural environment.

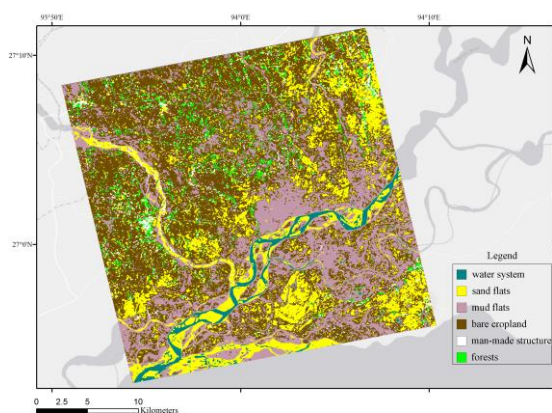


Figure 2. Sediment classification map on exposed intertidal flats using RF classifier based on full-polarimetric SAR features

	GROUP(i)	GROUP(ii)	GROUP(iii)
Overall accuracy	86.57%	82.14%	84.98%
Kappa coefficient	0.8464	0.7681	0.8014
Mud flats	85.12%	76.66%	83.32%
Sand flats	85.60%	74.85%	82.49%
Bare croplands	88.46%	80.61%	85.80%
forests	84.29%	86.84	89.58%

Table 2. Accuracy table of certain classes in each group

4. CONCLUSION AND DISCUSSION

Optical remote sensing data source can interpret vegetation and crops with higher accuracies; however, it performed lower capabilities of distinguishing different sediment types especially for mud flats, sand flats and bare croplands. But spectral information is complement with SAR indicators that can provide more detail information of targets with significant scattering disparities on intertidal flats. Radar sensors have great advantages on sediment classification. The surface roughness and soil moisture information can be obtained with full polarimetric SAR data especially in very heterogeneous and seasonally changing wetland. From variable importance indexes, we can see CP and FD components are promising when used to describe exposed sediments and soils, while Yamaguchi is not that suitable for sediments with similar radar backscattering signatures. Multiple data source acts with mix success for sediment discrimination. On the one hand, the asynchronous imaging time cause errors in rapidly changing targets; on the other hand, the optical channels can influence the contribution of SAR indicators for sediments with similar backscattering signatures.

Future work will still focus on proposing more SAR features for sediments on exposed intertidal flats, because small wetland areas are spread throughout the coastlines. They strongly need accurate and frequent monitoring because they are mainly threatened by human-induced activities that cause rapid processes of degradation with loss of wetland surface.

REFERENCES

- Baghdadi, N., Dubois-Fernandez, P., Dupuis, X., Zribi, M., 2013. Sensitivity of main polarimetric parameters of multifrequency polarimetric SAR data to soil moisture and surface roughness over bare agricultural soils. *Geoscience and Remote Sensing Letters*, IEEE 10, pp. 731-735.
- Banks, S.N., King, D.J., Merzouki, A., Duffe, J., 2014a. Assessing RADARSAT-2 for Mapping Shoreline Cleanup and Assessment Technique (SCAT) Classes in the Canadian Arctic. *Canadian Journal of Remote Sensing* 40, pp. 243-267.
- Banks, S.N., King, D.J., Merzouki, A., Duffe, J., 2014b. Characterizing Scattering Behaviour and Assessing Potential for Classification of Arctic Shore and Near-Shore Land Covers with Fine Quad-Pol RADARSAT-2 Data. *Canadian Journal of Remote Sensing* 40, pp. 291-314.
- Breiman, L., 2001. Random forests. *Machine learning* 45, pp. 5-32.
- Brockmann, C., Stelzer, K., 2008. Optical remote sensing of intertidal flats, *Remote Sensing of the European Seas*. Springer, pp. 117-128.
- Cloude, S.R., Pottier, E., 1996. A review of target decomposition theorems in radar polarimetry. *IEEE Trans. Geosci. Remote Sensing* 34, pp. 498-518.
- Corcoran, J., Knight, J., Brisco, B., Kaya, S., Cull, A., Murnaghan, K., 2012. The integration of optical, topographic, and radar data for wetland mapping in northern Minnesota. *Canadian Journal of Remote Sensing* 37, pp. 564-582.

Freeman, A., Durden, S.L., 1998. A three-component scattering model for polarimetric SAR data. *Geoscience and Remote Sensing, IEEE Transactions on* 36, pp. 963-973.

Gade, M., Alpers, W., Melsheimer, C., Tanck, G., 2008. Classification of sediments on exposed tidal flats in the German Bight using multi-frequency radar data. *Remote Sensing of Environment* 112, pp. 1603-1613.

Gade, M., Melchionna, S., Kemme, L., 2015. ANALYSES OF MULTI-YEAR SYNTHETIC APERTURE RADAR IMAGERY OF DRY-FALLEN INTERTIDAL FLATS. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Gade, M., Melchionna, S., Stelzer, K., Kohlus, J., 2014. Multi-frequency SAR data help improving the monitoring of intertidal flats on the German North Sea coast. *Estuarine, Coastal and Shelf Science* 140, pp. 32-42.

Koch, M., Schmid, T., Reyes, M., Gumuzzio, J., 2012. Evaluating Full Polarimetric C- and L-Band Data for Mapping Wetland Conditions in a Semi-Arid Environment in Central Spain. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 5, pp. 1033-1044.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67, pp. 93-104.

Ullmann, T., Schmitt, A., Roth, A., Duffe, J., Dech, S., Hubberten, H.-W., Baumhauer, R., 2014. Land Cover Characterization and Classification of Arctic Tundra Environments by Means of Polarized Synthetic Aperture X-and C-Band Radar (PolSAR) and Landsat 8 Multispectral Imagery—Richards Island, Canada. *Remote Sensing* 6, pp. 8565-8593.

van Beijma, S., Comber, A., Lamb, A., 2014. Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sensing of Environment* 149, pp. 118-129.