Ulva prolifera detection with dual-polarization GF-3 SAR data

Xiaomeng Geng¹, Pingxiang Li¹, Jie Yang¹, Lei Shi¹, Xiao-ming Li², Jinqi Zhao¹

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
2. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

Corresponding author: pxli@whu.edu.cn

Abstract: Ulva prolifera is a marine hazard occurs frequently in the Yellow Sea of China, which has led to continuous impacts on the marine environment. Accurate and timely information on Ulva prolifera monitoring by satellite remote sensing, plays a significant role in the marine ecological environment and coastal tourism development. Compared with optical remote sensing, Synthetic Aperture Radar (SAR) has the capability of imaging the earth's surface independent of sunlight and cloud, as well as in high spatial resolution. Polarimetric Synthetic Aperture Radar (PolSAR) backscattered signal that contains scattering and phase information, has a great advantage in target observation. An experiment based on GF-3 dual polarimetric characteristics features is implemented, which indicates it is suitable for detecting Ulva prolifera. The detection results are extracted based on the proposed method in Lianyungang coastal zone. Compared with the ground truth of expert interpretation, the kappa coefficient is 0.94, and the overall accuracy is 0.99.

Keywords: Ulva prolifera; SAR; GF-3; polarimetric characteristics

1. Introduction
In recent years, explosive proliferation of Ulva prolifera occurs frequently in the Yellow Sea of China due to global climate change and water eutrophication, which not only influences the balance of the ocean ecosystem but also brings serious threats to the development of coastal aquaculture and tourism. Huge efforts have been put to remove Ulva prolifera from these coastal seas in order to protect aquaculture and tourism. Therefore, monitoring the coverage of Ulva prolifera has become an imperative and practical issue. Field investigations by ship and airplane can understand the cause, however, detecting the process and state of Ulva bloom are not always practical due to the severe weather conditions or the difficulty of reaching offshore zones. Fortunately, the rapid development of satellite remote sensing technology has brought new observation methods for detecting the location and spatial coverage of Ulva prolifera. MODIS (Moderate Resolution Image Spectroradiometer), GOCI (Geostationary Ocean Color Imager), Landsat TM, HJ-1A/1B and FY-3A are widely used in Ulva prolifera detection. MODIS data covering the Yellow Sea zone twice a day that have been applied in Ulva bloom monitoring by the method of band ratio such as the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Floating Algae Index (FAI) [1-3]. However, the optical remote sensing is usually affected and not meet real-time application demand due
to the bad weather conditions and thick clouds. Whereas the Synthetic Aperture Radar (SAR) can continuously observe the earth with all day and night, and all-weather conditions. Therefore, SAR is priority to detect Ulva prolifera in cloudy and foggy covered areas. Currently, gray threshold and supervised classification such as likelihood classification, decision tree classification and random forest are used for detecting Ulva prolifera. Thresholding and clustering algorithms were used for detecting Ulva prolifera by the COSMO-SkyMed during the 2008 Olympic sailing competition[4,5]. Detection of Ulva prolifera was conducted for evaluation and results demonstrated that the random forest classification method had higher accuracy compared with the popular supervised classification methods (minimum distance and maximum likelihood) [6]. Moreover, most of these methods were based on single-polarized SAR by using gray features, texture, geometry, and other features[7,8]. However, the ocean surface backscattering coefficient is related to the physical mechanisms, which is not sufficient to investigate their characteristics or to improve their detection and classification for the complex ocean background. With the advent of dual-polarization and quad-polarization technology, more polarization shows wider potentially ability in classification and change applications than the single-polarized SAR mode. An index factor of co-pol SAR had been proved suitable for detecting green macroalgae[9]. Yamaguchi decomposition was sensitive to Enteromorpha prolifera for marine environmental monitoring of mixed floating sea-surface pollutants[10].

GF-3, the first Synthetic Aperture Radar (SAR) satellite of China, was successfully launched in August 2016, designed for land and ocean monitoring, disaster reduction and water resources evaluation and management. GF-3 system operates in twelve observation modes and supports single-polarization (HH or VV), dual-polarization (HH+HV or VH+VV) and quad-polarization (HH+HV+VH+VV) with the resolution varies from 1 m to 500 m, and the swath ranges from 10 km to 650 km[11]. Hence, this paper tends to provide a capability of detection of Ulva prolifera by using GF-3 SAR dual-polarization data. In Section 2, we introduce the study area, data selected and data preprocessed for this study. Section 3 demonstrates the two-component decomposition method which was applied for data processing in this study. In Section 4, the extracted polarimetric parameters are introduced for detecting Ulva prolifera. Finally, an optimum combination of polarimetric parameters is proposed. Conclusions are given in Section 5.

2. Study area and Dataset

2.1. Study area
The study area is located in the Yellow Sea of China, where is near the coastal zone of Lianyungang city (See Figure 1a), with the longitude and latitude of 34°17′-35°26′, 119.76°-121°15′ respectively. Annually, large-scale blooms of Ulva prolifera usually occurs in the middle of the Yellow Sea from May to June [12], and then drifting to the offshore areas of Shandong province due to the wind and sea currents, which bring about severe threats to ocean ecosystems.
Figure 1. The image of GF-3 dual-polarization SAR captured in July 20, 2019: (a) the location of google earth; (b) HH polarization (uint: dB); (c) HV polarization (uint: dB); (d) and (e) are photos of Ulva prolifera

2.2. Dataset
In this study, the GF-3 dual-polarization SAR image was collected on June 20, 2019, over China’s Yellow Sea covered by vast Ulva bloom. Besides Single Look Complex (SLC), we used Fine Stripe Mode II (FSII) dual-polarization (HH, HV) with the spatial coverage of 100 km×100 km spatial coverage and resolution of 10m. Detailed information of the image is listed in Table 1. The corresponding scattering amplitude and photos are shown in Figure 1b to Figure 1e.

Table 1. The main parameters of the C-band GF-3 dual-polarization SAR image

| Location          | Time        | Acquisition mode | Incidence angle | Polarization | Resolution | Swath    |
|-------------------|-------------|------------------|-----------------|--------------|------------|----------|
| The Yellow Sea of China | 2019-6-20 10:09 | FSII            | 46.52°~51.34°  | HH, HV       | 10 m       | 100 km   |

2.3. Data preprocessing
The GF-3 PolSAR data is pre-processed, including radiometrically calibration, Sinclair scattering matrix extraction, second-order coherence covariance matrix T2 extraction, $H - \alpha$ decomposition and polarimetric characteristics feature extraction. Multilook process (4×4) and sigma Lee filtering (7×7) were applied to reduce speckle filter influences when pre-processed. A subset of 2000×4000 was clipped after multi-look in order to save time and verificate algorithm. The ground truth data we used in the study is obtained by expert interpretation. Figure 2 shows the data preprocess flowchart.
3. Method

HH and HV dual-polarization information can be expressed by Sinclair scattering matrix and the covariance matrix C2, which is similar to full-polarized SAR.

The HH-HV SAR scattering matrix is defined in equation (1):

\[
S = \begin{bmatrix} S_{HH} & S_{HV} \\ 0 & 0 \end{bmatrix}
\]

The corresponding scattering vector is defined in equation (2):

\[
k = \left[ S_{HH} \frac{S_{VV} + jS_{HV}}{\sqrt{2}} \right]^T
\]

\[
C_{HHHV} = k \cdot k^* \begin{bmatrix} \langle |S_{HH}|^2 \rangle & \langle S_{HH}S_{HV}^* \rangle \\ \langle S_{HV}S_{HH}^* \rangle & \langle |S_{HV}|^2 \rangle \end{bmatrix} = U \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} U^T = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T
\]

Where, \( U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \) and \( u_i = e^{i\theta} \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix} \). \( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues of \( C_{HHHV} \).

The superscript \( T \) denotes the conjugate transpose and \( * \) denotes the conjugate.

The polarimetric entropy \( H \), scattering angle \( \alpha \) and anisotropy \( A \) are defined as follows:

\[
H = \sum_{i=1}^{2} -P_i \log_2 P_i
\]

\[
\alpha = \sum_{i=1}^{2} P_i \cos^{-1}(u_i)
\]

\[
A = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}
\]

The HA, \( H(1-A) \), \( 1-H \), and \( (1-H)(1-A) \) can be obtained by combine between entropy \( H \) and anisotropy \( A \).

The pseudo-probabilities can be defined from the set of sorted eigenvalues in equation (7):

\[
P_i = \frac{\lambda_i}{\sum_{i=1}^{2} \lambda_i}, i = 1, 2
\]

The Shannon entropy (SE) is defined as the sum of two components related to intensity and polarization.
\[
SE = \log(\pi^3 e^3 |T_2|) = SE_t + SE_p
\]  
(8)

where, \(SE_t\) represents the intensity contribution related to total span; \(SE_p\) represents the polarimetric contribution related to the degree of polarization.

\[
SE_t = 3\log(\pi e\text{Tr}(T_2))
\]

\[
SE_p = \log(1 - (1 - 27 \frac{|T_2|}{\text{Tr}(T_2)^3})^{1/2})
\]  
(9)

Other polarimetric features can be found in [13]. Due to the space limitation, we do not discuss them in detail here.

4. Results and Discussion

4.1. Polarimetric characteristics of Ulva prolifera

A subset of size 2000×4000 was cropped after the GF3 multilook in order to analysis the polarimetric characteristics of Ulva prolifera easily and quickly. Figure 3 shows the amplitude of the normalized backscattered radar signals in HH and HV polarizations. The open seawater in HH and HV polarization images was dark due to the smooth sea surface, which is dominant by surface scattering. In contrast, Ulva prolifera appears as a bright target in HH and HV polarization images. Since Ulva prolifera is a kind of multicellular organism in some species form filamentous with an erect much-branched or single-branch frond, its specific structures could have a contribution to radar backscattered signals at both polarizations. Hence, the volume scattering is the dominant scattering mechanism in HV, whereas surface or Bragg scattering mechanism is more important in HH polarization. One should note the HV-polarized signals backscattered from Ulva prolifera and sea surface are usually quite weak and clean than HH polarization due to low signal-to-noise ratio. In addition, it is noticeable that even though after Lee filtering, there are filamentous features in the upper right corner of the HH polarization (see Figure 4 black rectangle), which is not seen in HV polarization. The backscattering coefficient of irregular pattern is close to Ulva prolifera (name as the competing signals or noise), which makes Ulva prolifera more difficult to identify in SAR images. Although many studies about the detection of Ulva prolifera had been conducted in the previous study[4-6,14], this special phenomenon is seldom reported. The competing signals or noise in HH polarization might be related to other atmospheric and ocean dynamics[9], or low pulse repetition frequency. As mentioned before, HV is dominated by volume scattering instead of Bragg scattering. In this case, assuming it is related to the ocean and atmospheric dynamics, the process only modulates Bragg waves, thereby making the irregular patterns present in HH. Moreover, it is possible to be related to the azimuth ambiguity that appears as ghost in SAR images due to low pulse repetition frequency. However, the reliable reason why the noise only exists in HH polarization need further analysis and verification.
4.2. Dual-polarization $H$-$\alpha$ decomposition

Fortunately, since GF3 data is received in dual-polarization SLC mode, two-component $H$-$\alpha$ decomposition can be developed to identify *Ulva prolifera* from the open seawater. The alpha and entropy are firstly decomposed based on HH and HV as shown in Figure 5a and 5b. The scattering mechanism of open seawater and *Ulva prolifera* shows not easy to distinguish in the alpha and entropy. In addition, *Ulva prolifera* and competing signals are more obvious with lower values in alpha and entropy compared with the open seawater area. Our main goal is to differentiate *Ulva prolifera* from open seawater. Hence, 1000 sample points were selected to demonstrate whether *Ulva prolifera* and open seawater can be discriminated or not. In Figure 5c, *Ulva prolifera* shows a lower entropy and alpha than open seawater area, which range from 0.56 to 0.82 and 13.44° to 24.33°, respectively. And the seawater shows similar results to [10], which range from 0.71 to 0.89 in entropy and 18.61° to 28.75° in alpha, respectively. This means it is quite difficult to distinguish *Ulva prolifera* and open seawater only referring to alpha and entropy. Similarly, Dual-polarization $H$-$\alpha$ decomposition was employed to discuss the scattering mechanisms and classify the land covered type [15], in which the $H$-$\alpha$ plane, HH-HV and HV-VV cannot effectively illustrate the scattering mechanisms and land covers as well.
4.3. Polarimetric features parameters based on dual-polarization

In qual-polarization SAR, the target classification and detection can be improved by combining polarimetric alpha, entropy, anisotropy and other polarimetric parameters. In dual-polarization SAR, the scattering mechanism and polarimetric features have been discussed in land classification, but less discussed in Ulva prolifera detection. Therefore, in order to uncover the most powerful parameters for detecting Ulva prolifera, a series of polarimetric parameters were extracted by combining HH and HV. All the polarimetric features parameters were extracted and processed by European Space Agent (ESA) PolSARPro Software. In those parameters, $f_1$, $f_2$, $f_3$, $f_4$ and $f_5$ are extracted by combining entropy and anisotropy, which are defined as CMBA in Table 2. $f_6$, $f_7$, $f_8$, $f_9$, $f_{10}$ and $f_{11}$ correspond to Shannon entropy and component parameters (SECP). $f_{12}$ and $f_{13}$ correspond to pseudo-probabilities (PP). $f_{14}$, $f_{15}$ and $f_{16}$ are eigenvalues and lambda (EL). $f_{17}$ and $f_{18}$ are the covariance matrix (CM). Table 2 lists all feature parameters, which are explained in detail in [13]. Figure 6 shows the extracted feature parameter results.

In those features, $f_1$, $f_2$, $f_3$, $f_4$, $f_5$, $f_6$, $f_{10}$, $f_{11}$, $f_{12}$ and $f_{13}$ fail to distinguish Ulva prolifera, but with some interesting phenomenon that deserves attention in Figure 6. Among those, $f_6$ and $f_9$ show the value of Ulva prolifera higher than open seawater and competing noise, but the value of open seawater is high in near-range direction and easily gets confused with Ulva prolifera. In Figure 3 we can also find the calibration in near-range is more difficult in HH, which might affect the values of $f_6$ and $f_9$ in open seawater. In addition, $f_3$, $f_6$, $f_7$, $f_{14}$, $f_{15}$, $f_{16}$, $f_{17}$ and $f_{18}$ can distinguish Ulva prolifera from open seawater (See Figure 6). The competing noise can be suppressed in $f_{15}$ and $f_{18}$ because those parameters are mainly based on cross-polarization. The scattering mechanism of targets is more physical significance for qual-polarization, and the performance of HH-VV SAR for extracting scattering mechanisms more effectively than HH-HV and HV-VV has been proved theoretically [15]. Therefore, we do not discuss the scattering mechanism of Ulva prolifera further but focus on the most sensitive features for distinguishing Ulva prolifera from open seawater.

4.4. The feature importance scores of dual-polarization and Random Forest (RF) classification

In the above section, the polarimetric characteristics of dual-polarization are discussed. In this section, we mainly concern which dual-polarization feature is more suitable for detecting Ulva prolifera. Therefore, RF is introduced in order to determinate which features are most suitable for detecting Ulva prolifera. RF [16] is machine learning algorithm that builds multiple trees based on random bootstrapped samples of the training data, which is robust to overfitting and can handle thousands of input variables. RF can not only classify land cover but also estimate the importance of each predictor variable in the model [17]. In this experiment, about 1000 samples of each class (Ulva prolifera and open seawater) were selected. We use about 80% of selected data to construct the classification tree and validated this classifier on the remaining 20% of the data. Random split variables at each node is 4 and the number of trees is set to 200 for identifying the importance score of each feature.

Figure 7 shows the importance score for each the polarimetric features derived above mentioned. Apparently, $f_7$, $f_{14}$, $f_{15}$ and $f_{16}$ has a higher importance score than the other. The results show Ulva prolifera...
prolifera can be better distinguished from seawater in Figure 8. Moreover, the signal of Ulva prolifera

| Table 2. Polarization features extracted with GF3 dual-polarization SAR |
|-------------------|------------------|
| Index     | Feature parameters                                      |
| CMBA      | anisotropy \((f_1)\), (1-H)(1-A) \((f_2)\), HA \((f_3)\), (1-HA) \((f_4)\), H(1-A) \((f_5)\) |
| SECP      | entropy_shannon \((f_6)\), entropy_shannon_I \((f_7)\), entropy_shannon_I_norm \((f_8)\), entropy_shannon_norm \((f_9)\), entropy_shannon_P \((f_{10})\), entropy_shannon_P_norm \((f_{11})\) |
| PP        |                                                 |
| EL        |                                                 |
| CM        |                                                 |

Figure 6. Results of GF-3 feature parameters: \((f_1) \sim (f_{18})\) represent the listed features in Table 2.

in \(f_7\), \(f_{14}\) and \(f_{16}\) is obvious than \(f_{15}\), whereas the competing noise is better suppressed in \(f_{15}\). Therefore, we select four features with higher importance scores to classify Ulva prolifera based on RF model. The RF classifier can provide probabilistic classification, which can be used to obtain the corresponding binary classification result by setting a threshold. Since the results are highly correlated with the threshold set, the influence of a threshold used to produce the binary classification is studied. Figure 9a shows the default threshold value of 0.5 will give 0.994 overall accuracy and 0.92 kappa coefficient, and the best results can be obtained with the threshold value of 0.75. Hence, Ulva prolifera is detected by threshold value of 0.75, and the result is shown in Figure 9b. The accuracy is 0.99 and the kappa coefficient is 0.97 by ground truth as expert interpretation (Figure 9c).

5. Conclusions
In this paper, we studied the polarimetric characteristics of Ulva prolifera in GF3 dual-polarization HH
Figure 7. The importance scores of the polarimetric features: 1~2 represents alpha and entropy, and 3~20 represent f1~f18 in Table.2.

Figure 8. The important features: (a) f15; (b) f7; (c) f14; and (d) f16.

Figure 9. The results of Ulva prolifera classification: (a) the binary map of Ulva prolifera (dark blue: Ulva prolifera, white: seawater); (b) the accuracy of the different threshold value; (c) the ground truth of expert interpretation.

and HV data. The backscattering coefficient of Ulva prolifera was obviously stronger than the open seawater in both polarization, whereas, HH was easily influenced by competing signals and unavailable for detecting Ulva prolifera. H-\(\alpha\) two-decomposition and more polarimetric features were taken to discuss the importance scores of Ulva prolifera detection. The importance scores of features demonstrate f15, f7, f14 and f16 were suitable for detecting Ulva prolifera, respectively. The classification accuracy and kappa coefficient of Ulva prolifera were 0.99 and 0.97, respectively, when
the threshold was set 0.75. It was demonstrated that detecting *Ulva prolifera* was available and could obtain better results by dual-polarization SAR.

Comparing with qual-polarization data, dual-polarization (HH or VV, HH or HV, HV or VV) has large swath widths to satisfy the requirement of wide-ocean-area monitoring but less researched. Hence, in our future study, we will mainly concern and discuss the spatial-temporal distribution of *Ulva prolifera* in coastal China.

**Acknowledgments**

This work is supported in part by the National Natural Science Foundation of China under Grant 41601355, Grant 41771377. The author would like to National Satellite Ocean Application Service for providing the GF3 SAR image, and the reviewers for the suggestions that markedly improved this paper.

**Reference**

[1] Wu S Shao F Wang Y Sun R and Wang J 2014 Enteromorpha Prolifera Detection with MODIS Image Using Semi-supervised Clustering *JCP* 9 1259-1265.

[2] Hu L Zeng K Hu C and He M-X 2019 On the remote estimation of Ulva prolifera areal coverage and biomass *Remote Sens. Environ.* 223 194-207.

[3] Xiao Y Zhang J Cui T Gong J Liu R Chen X and Liang X 2019 Remote sensing estimation of the biomass of floating Ulva prolifera and analysis of the main factors driving the interannual variability of the biomass in the Yellow Sea *Mar. Pollut. Bull.* 140 330-340.

[4] Wu D Zhang B Li J Wu Y Zhang H and Shen Q 2009 Monitoring of Enteromorpha prolifera in Qingdao marine by exploiting the synergy of active and passive remote sensing data *IGARSS* I-228-I-231.

[5] Wang S Zhang F Shao Y Tian W and Gong H 2010 Microwave remote sensing for marine monitoring: An example of Enteromorpha prolifera bloom monitoring *IGARSS* 4530-4533.

[6] Xie C Dong J Sun F and Bing L 2016 Object-Oriented Random Forest Classification for Enteromorpha Prolifera Detection with SAR Images *ICVRV* 119-125.

[7] Hu C 2009 A novel ocean color index to detect floating algae in the global oceans *Remote Sens. Environ.* 113 2118-2129.

[8] Wang X Li L Bao X and Zhao L 2009 Economic cost of an algae bloom cleanup in China's 2008 Olympic sailing venue *Eos, Transactions American Geophysical Union* 90 238-239.

[9] Shen H Perrie W Liu Q and He Y 2014 Detection of macroalgae blooms by complex SAR imagery *Mar. Pollut. Bull.* 78 190-195.

[10] Wang X Shao Y Tian W and Li K 2018 On the classification of mixed floating pollutants on the Yellow Sea of China by using a quad-polarized SAR image *Front. Earth Sci* 12 373-380.

[11] Sun J W Y and Y D 2017 The SAR Payload Design and Performance for the GF-3 Mission *Sensors* 17 2419.

[12] Liu X Li Y Wang Z Zhang Q and Cai X 2015 Cruise observation of Ulva prolifera bloom in the southern Yellow Sea, China *Estuar. Coast. Shelf. Sci.* 163 17-22.

[13] Lee J-S and Pottier E 2009 Polarimetric radar imaging: from basics to applications. CRC press.

[14] Li Y Liang G Yu S and Chen P 2011 Selection of microwave remote sensing data of monitoring of Enteromorphaprolifera disaster *Marine Environmental Science* 30 739-742.

[15] Ji K and Wu Y 2015 Scattering mechanism extraction by a modified cloude-pottier decomposition for dual polarization SAR *Remote Sens.* 7 7447-7470.

[16] Breiman L 2001 Random Forests *Mach. Learn.* 45 5-32.

[17] Genuer R Michel J Christine P and Malot T 2010 Variable selection using random forests *Pattern Recognit. Lett.* 31 2225-2236.