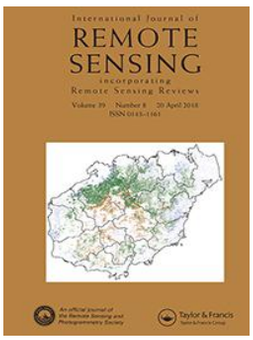

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

# MSMD: maximum separability and minimum dependency feature selection for cropland classification from optical and radar data

Iman Khosravi, Abdolreza Safari & Saeid Homayouni

**To cite this article:** Iman Khosravi, Abdolreza Safari & Saeid Homayouni (2018) MSMD: maximum separability and minimum dependency feature selection for cropland classification from optical and radar data, International Journal of Remote Sensing, 39:8, 2159-2176, DOI: 10.1080/01431161.2018.1425564

**To link to this article:** https://doi.org/10.1080/01431161.2018.1425564

INTERNATIONAL JOURNAL OF REMOTE SENSING, 2018
VOL. 39, NO. 8, 2159–2176
https://doi.org/10.1080/01431161.2018.1425564

Taylor & Francis
Taylor & Francis Group



# MSMD: maximum separability and minimum dependency feature selection for cropland classification from optical and radar data

Iman Khosravi[a], Abdolreza Safari[a] and Saeid Homayouni[b]

[a]School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran; [b]Department of Geography, Environment, and Geomatics, University of Ottawa, Ottawa, Canada

**ABSTRACT**

Cropland classification using optical and full polarimetric synthetic aperture radar (PolSAR) images is a topic of considerable interest in the remote-sensing community. These two data sources can provide a diverse set of temporal, spectral, textural and polarimetric features which can be invaluable for cropland classification. However, some optical features or some radar features may have a relatively high correlation with other features. Hence, it seems to be necessary to choose the optimum features in order to reduce the dimensions of the data and to improve cropland classification accuracy. This article proposes a strategic feature selection method from a feature set of bitemporal RapidEye and Uninhabited Aerial Vehicle synthetic aperture radar (UAVSAR) images. The proposed method is designed to select the most relevant features and to remove redundant features based on the two concepts of separability and dependency. The proposed method is therefore referred to as maximum separability and minimum dependency (MSMD). For evaluating efficiency, MSMD and some well-known filter and wrapper feature selection methods are compared using a random forest classifier. Experimental tests confirmed that the classification results obtained from the MSMD feature selection method were more accurate than those achieved by filter methods. Moreover, they had an accuracy comparable to that of the results from the wrapper method. Furthermore, with regard to running time, MSMD operated as fast as the filter methods. It had a straightforward structure compared to the wrapper method, and as a result was faster than this method.

**ARTICLE HISTORY**
Received 12 July 2017
Accepted 28 December 2017

## 1. Introduction

Cropland classification is an essential step in agricultural planning and management activities at national and global scales (Wardlow and Egbert 2010). The crop map plays an important role in various applications such as crop inventories, crop insurance, yield estimation and the enforcement of quota limits (Foody, Mcculloch, and Yates 1994). Earth observation technology makes use of optical and radar sensors with various spatial and temporal resolutions. Observations from these sensors can

**CONTACT** Abdolreza Safari asafari@ut.ac.ir School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

be used for agricultural land classification in a more efficient manner and with lower costs than traditional survey methods (Xie et al. 2007).

Numerous studies to date have utilized Earth observations from optical sensors for classification of croplands (Joshi et al. 2016). The most widely used optical sensors are the Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat, NOAA and IRS sensors, due to their higher revisit time. RapidEye has been used in several recent studies (e.g. Löw et al. 2012; Kim and Yeom 2014; Siachalou, Mallinis, and Tsakiri-Strati 2015; Kross et al. 2015). A unique spectral band, the so-called red edge (RE), has been embedded in the RapidEye sensor in order to analyse vegetation. Multispectral bands, vegetation indices (VIs) and textural indicators are the main features extracted from the optical sensors which have been used for crop mapping. These features provide valuable information about the vegetation reflectance in the visible and infrared regions of electromagnetic spectrum, or the spatial characteristics of the crop types (De Souza et al. 2015; Peña-Barragán et al. 2016).

Compared to those using optical sensors, fewer studies have used synthetic aperture radar (SAR) or full polarimetric SAR (PolSAR) observations for crop mapping (Joshi et al. 2016). This may be due to the higher cost and lower availability of PolSAR data. RADARSAT is the most widely used PolSAR sensor for crop mapping. In addition, the Uninhabited Aerial Vehicle synthetic aperture radar (UAVSAR) sensor has been recently used in several studies for agricultural applications (e.g. Mahdian et al. 2013; Yekkehkhany et al. 2014a; Yekkehkhany et al. 2014b; Hosseini et al. 2015; Khosravi et al. 2017; Tamiminia et al. 2017; Whelen and Siqueira 2017). UAVSAR has an L-band frequency and a tolerable revisit time; features which are suitable for analysing crop types. Various parameters from the PolSAR sensors can be used for cropland classification. These are backscattering intensities, polarization discriminators and scattering parameters extracted from target decompositions (Alberga 2007; Alberga et al. 2008). These features give information about the physical and structural properties and also the scattering mechanisms of the various crop types (Hoang et al. 2016).

In the past two decades, several researchers have used a combination or fusion of optical and SAR/PolSAR images for cropland classification (Joshi et al. 2016). The joint use of optical and radar data has two advantages. First, it helps to mitigate the specific limitations of each type, i.e. dependency on weather and atmospheric conditions in optical images and speckle defects in radar images. Second, it takes advantage of their complementarity regarding sensitivity to vegetation and soil characteristics (Soria-Ruiz, Fernandez-Ordonez, and Woodhouse 2010; Villa et al. 2015). For example, Kussul et al. (2013) compared six optical images from MODIS, TM, ALI, AWiFS, LISS-3, and RapidEye for crop mapping, and found that the use of optical data alone led to limitations in detecting some summer crop types in Ukraine. Therefore, they used C-band RADARSAT-2 data to improve the mapping accuracy. Forkuor et al. (2014) used a combination of RapidEye and TerraSAR-X images and were able to improve overall crop-mapping accuracy by up to 10–15%. In a recent paper, data from two novel optical and radar satellites, i.e. Landsat-8 and Sentinel-1, were used for the classification of agricultural fields (Inglada et al. 2016). The results indicated a considerable increase in accuracy compared to the individual data sources.

On the other hand, the characteristics of crops change during the growing stages. Several studies have addressed the fact that the temporal combination of optical and

PolSAR observations can provide additional valuable information for crop mapping (e.g. Zhu and Tateishi 2006; McNairn et al. 2009; Haldar and Patnaik 2010; Torbick et al. 2011; Villa et al. 2015; Kussul et al. 2016; Park and Im 2016; Skakun et al. 2016). Accordingly, this combination could generate a full and large optical-radar data set with an exhaustive temporal, spectral, textural and polarimetric feature space for cropland classification. However, there may be relatively high correlations among some optical features or some polarimetric features. Moreover, some features may have low relevancy with respect to some or all classes. Therefore, it is necessary to choose the best features for cropland classification, which have the most relevancy with respect to the classes and the lowest redundancy among themselves.

There are, in general, two main types of feature selection algorithms in the literature: filter and wrapper methods. Filter methods can operate very fast; however, they solely may not always lead to a significant increase in accuracy. By contrast, wrapper methods run very slowly. Moreover, they are usually limited to their respective classifiers and may not improve the accuracy when using other classifiers (Zeng et al. 2015). Therefore, in this article, we propose a strategic, intrinsic feature selection procedure from a feature set consisting of optical and PolSAR data, for crop mapping. The study area is an agricultural region near Winnipeg, Manitoba, Canada. The data are bitemporal RapidEye and UAVSAR images. The proposed algorithm employs the two concepts of separability and correlation for selecting the optimum features. For that reason, we name it the maximum separability and minimum dependency (MSMD) method. The accuracy and running time of the proposed method are compared with those of some well-known filter and wrapper feature selection algorithms using the random forest (RF) classifier. This method could be advantageous for operational crop-mapping activities from freely available Earth observations.

## 2. Earth observation data

### 2.1. *Study area and remote-sensing data*

The study area was the south-west district of Winnipeg, Manitoba, Canada, which is covered with various agricultural crops (see Figure 1). It was the study site for the Soil Moisture Active Passive Mission (SMAP) Validation Experiment 2012 (SMAPVEX12) campaign for the calibration and validation of NASA's SMAP satellite (McNairn et al. 2015). The region used in this study is located between latitudes 47° 32′ 16″ N and 48° 12′ 56″ N, and between longitudes 97° 5′ 2″ W and 97° 45′ 13″ W. The area of the region is approximately 400 ha. The mean elevation from sea level is around 784 ft (239 m).

The data used in this article were bitemporal observations from RapidEye and UAVSAR sensors acquired on 5 and 14 July 2012 (see Table 1). The RapidEye images had five spectral channels including blue (B), green (G), red (R), near-infrared (NIR) and red edge (RE). They were orthorectified on the local North American 1983 datum (NAD-83) and had a spatial resolution of about 5 m. The UAVSAR images had quad polarizations of HH, HV, VH, and VV acquired at the L-band frequency. 'H' and 'V' indicate horizontal and vertical directions, respectively. These PolSAR images were orthorectified on the World Geodetic System 1984 datum (WGS-84) with an SRTM3 digital elevation model. They were also multi looked by two pixels in the azimuth direction and three pixels in the range direction.

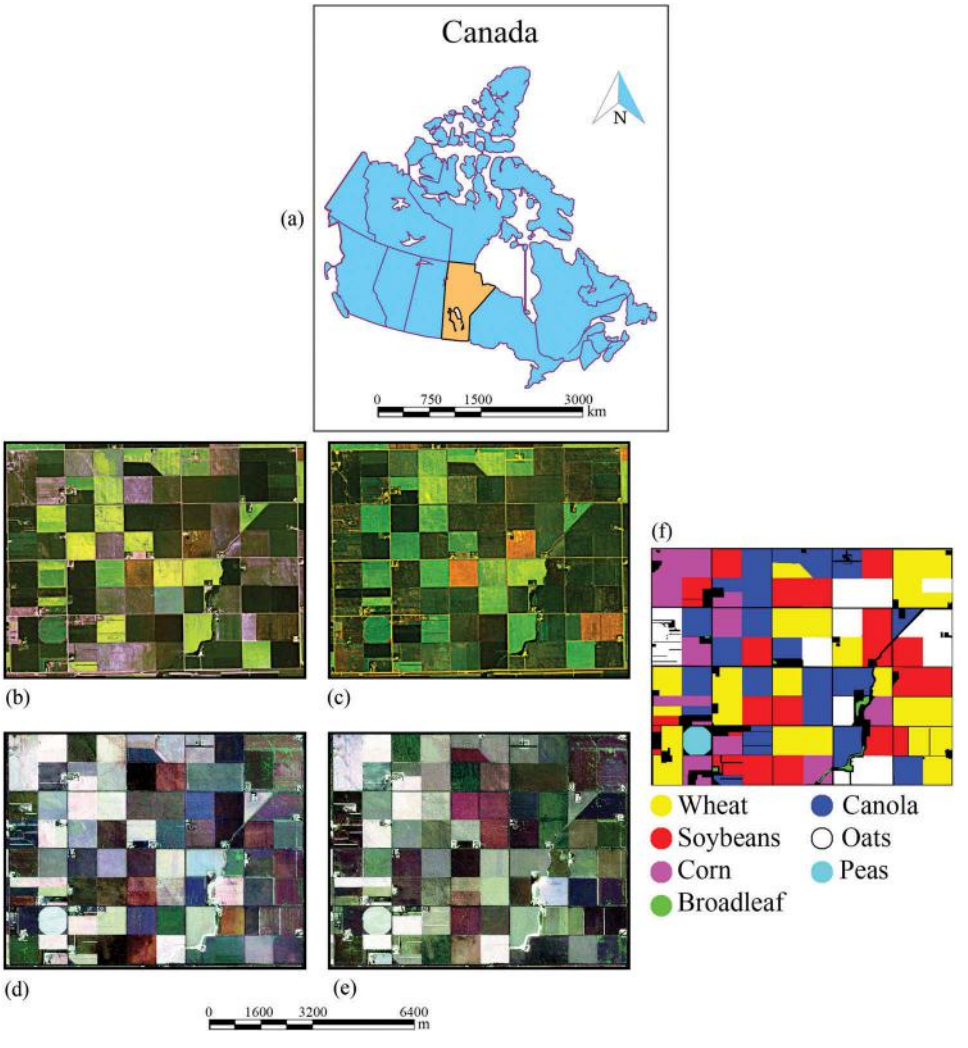


**Figure 1.** The optical and PolSAR images used in this article from (a) Winnipeg, Manitoba, Canada: (b) and (c) show the true color composite of optical images acquired on 5 and 14 July 2012, (d), and (e) show the Pauli RGB colour composite of the PolSAR images acquired on 5 and 14 July 2012. (f) The reference data of the study area.

**Table 1.** Properties of the optical and PolSAR images used in this article.

| Sensor | Channels | Resolution (m) | Size (pixels) |
|---|---|---|---|
| RapidEye (Satellite) | Blue (B), green (G), red (R), red-edge (RE), and near-infrared (NIR) spectral bands | 5 × 5 | 1412 × 1711 |
| UAVSAR (Airborne) | L-band HH, HV, VH, and VV | 15 × 15 | 551 × 701 |

Moreover, a despeckling process using a 5 × 5 boxcar filter (based on trial and error experiments), was applied to the data in order to reduce the speckle effect. The spatial resolution of these images was then approximately 15 m. Finally, the bitemporal optical and PolSAR images were efficiently co-registered.

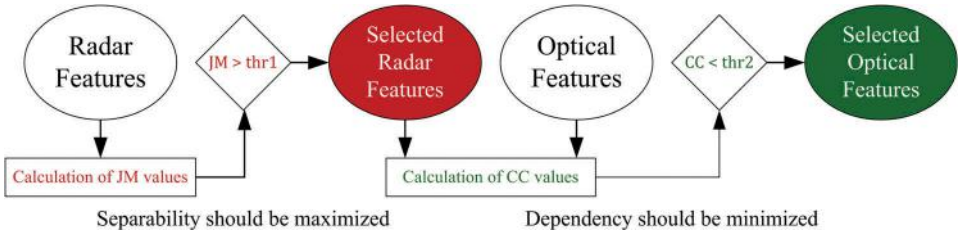


**Figure 2.** The procedure of MSMD feature selection algorithm.

The reference data were collected during the SMAPVEX12 campaign which took place over 43 days during the summer of 2012. According to the reference data for the study area (see Figure 2(e)), seven classes were found in this region: wheat, canola, soybeans, oats, corn, peas and broadleaf. The number of samples for each class is shown in Table 2. In this study, only 5% of all samples were randomly selected as training samples to build the classifier model, and the remaining samples (i.e. 95%) were used for testing and evaluating the classifier.

## 2.2. *Extracting features*

### 2.2.1. *Extracting optical features*

Five spectral bands formed the primary features of the RapidEye images. Furthermore, a variety of common VIs, including a set of classical VIs based on R (Peña-Barragán et al. 2016) and a set of VIs based on RE (Kross et al. 2015), as presented in Table 3, were extracted from the optical images.

In addition to spectral and VI features, textural indicators were extracted from the RapidEye images. In this article, the eight main parameters of the grey-level co-occurrence (GLCM) matrix, i.e. mean ($\boldsymbol{\mu}$), variance ($\boldsymbol{\sigma}$), homogeneity (HOM), contrast (CON), dissimilarity (DIS), entropy ($\boldsymbol{H}$), angular second moment (ASM) and correlation (COR), were used to describe the texture (Rao et al. 2002; Theodoridis and Koutroumbas 2003). Kim and Yeom (2014) reported that the GLCM of the RE and NIR bands can be useful in improving crop-mapping accuracy. However, in this study, we adopted the following procedure: we initially applied a principal component analysis (PCA) to the spectral channels and then chose the PCs which had the highest information content. Based on the experiment, we used PC1 and PC2 for both the RapidEye images because they contained 99.03% and 98.35%, respectively, of all information in all channels. Subsequently, GLCM was applied to these two features rather than to the spectral channels. According to our trials, these textural

**Table 2.** Number of all samples for the study area.

| Class | Total number of samples | Number of training samples | Number of test samples |
|---|---|---|---|
| Wheat (Wh) | 85,074 | 4250 | 80,824 |
| Canola (Ca) | 75,673 | 3780 | 71,893 |
| Soybeans (So) | 74,067 | 3700 | 70,367 |
| Oats (Oa) | 47,117 | 2360 | 44,757 |
| Corn (Co) | 39,162 | 1960 | 37,202 |
| Peas (Pe) | 3598 | 180 | 3418 |
| Broadleaf (Br) | 1143 | 55 | 1088 |

**Table 3.** Optical features extracted from two RapidEye images.

| Spectral bands | B: 440–510 nm, G: 520–590 nm, R: 630–685nm, RE: 690–730 nm, NIR: 760–850 nm | |
|---|---|---|
| Vegetation indices based on red | Normalized difference vegetation index | $NDVI = (NIR - R)/(NIR + R)$ |
| | Simple ratio | $SR = NIR/R$ |
| | Red-green ratio index | $RGRI = G/R$ |
| | Enhanced vegetation index | $EVI = 2.5(NIR - R)/(NIR + 6R - 7.5B + 1)$ |
| | Atmospherically resistant vegetation index | $ARVI = (NIR - (2R - B))/(NIR + (2R - B))$ |
| | Soil-adjusted vegetation index | $SAVI = 1.5(NIR - R)/(NIR + R + 0.5)$ |
| | Normalized difference greenness index | $NDGI = (G - R)/(G + R)$ |
| | Green NDVI | $gNDVI = (NIR - G)/(NIR + G)$ |
| | Modified triangular vegetation index | $MTVI2 = \frac{1.5(1.2(NIR - G) - 2.5(R - G))/}{\sqrt{(2NIR + 1)^2 - (6NIR - 5\sqrt{R}) - 0.5}}$ |
| Vegetation indices based on red-edge | Red edge normalized difference vegetation index | $NDVIre = (NIR - RE)/(NIR + RE)$ |
| | Red edge simple ratio | $SRre = NIR/RE$ |
| | Red edge normalized difference greenness index | $NDGIre = (G - RE)/(G + RE)$ |
| | Red-edge triangular vegetation index | $RTVIcore = 100(NIR - RE) - 10(NIR - G)$ |
| | Red edge NDVI | $RNDVI = (RE - R)/(RE + R)$ |
| | Transformed CARI | $TCARI = 3((RE - R) - 0.2(RE - G)(RE/R))$ |
| | Triangular vegetation index | $TVI = 0.5(120(RE - G) - 200(R - G))$ |
| | Red edge ratio 2 | $PRI2 = RE/R$ |
| Textural indicators | $\boldsymbol{\mu}_{PC1}$, $\boldsymbol{\sigma}_{PC1}$, $HOM_{PC1}$, $CON_{PC1}$, $DIS_{PC1}$, $\boldsymbol{H}_{PC1}$, $ASM_{PC1}$, $COR_{PC1}$ | |
| | $\boldsymbol{\mu}_{PC2}$, $\boldsymbol{\sigma}_{PC2}$, $HOM_{PC2}$, $CON_{PC2}$, $DIS_{PC2}$, $\boldsymbol{H}_{PC2}$, $ASM_{PC2}$, $COR_{PC2}$ | |

features could give higher accuracy (approximately 5% higher) than the textural features obtained by RE and NIR bands. These 38 optical features are presented in Table 3.

### 2.2.2. *Extracting radar features*

The primary features we extracted from both the UAVSAR images were three linear and three circular polarization intensities ($\boldsymbol{\sigma}_{HH}$, $\boldsymbol{\sigma}_{HV}$, and $\boldsymbol{\sigma}_{VV}$; $\boldsymbol{\sigma}_{RR}$, $\boldsymbol{\sigma}_{RL}$, and $\boldsymbol{\sigma}_{LL}$). In addition, three ratios of linear polarizations ($\boldsymbol{R}_{HHVV}$, $\boldsymbol{R}_{HVHH}$, and $\boldsymbol{R}_{HVVV}$) and three ratios of circular polarizations ($\boldsymbol{R}_{RRLL}$, $\boldsymbol{R}_{RLRR}$, and $\boldsymbol{R}_{RLLL}$) were extracted. Three ratios of linear polarizations to total power, i.e. span ($\boldsymbol{R}_{HH}$, $\boldsymbol{R}_{HV}$, and $\boldsymbol{R}_{VV}$) and three ratios of circular polarizations to span ($\boldsymbol{R}_{RR}$, $\boldsymbol{R}_{RL}$, and $\boldsymbol{R}_{LL}$) were also extracted. Moreover, correlation coefficients between pairwise polarizations ($\boldsymbol{\rho}_{HHVV}$, $\boldsymbol{\rho}_{HVHH}$, and $\boldsymbol{\rho}_{HVVV}$; $\boldsymbol{\rho}_{RRLL}$, $\boldsymbol{\rho}_{RLRR}$, and $\boldsymbol{\rho}_{RLLL}$), due to the different values of various land-cover types, were extracted from both the UAVSAR images (Lardeux et al. 2009; Lönnqvist et al. 2010; Entezari, Motagh, and Mansouri 2012; Yekkehkhany et al. 2014a; Chen, Li, and Wang 2015). Here, 'H' and 'V' denote horizontal and vertical linear polarization, and 'R' and 'L' denote right-hand and left-hand circular polarization.

Other radar features extracted from the PolSAR data were statistical features related to the eigenvalues of the coherency matrix. These features were: three eigenvalue decompositions ($\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, and $\boldsymbol{\lambda}_3$) representing the weight indices of each scattering mechanism, the three main parameters of entropy ($\boldsymbol{H}$), anisotropy ($\boldsymbol{A}$), and alpha ($\bar{\alpha}$) together with four combinations of $\boldsymbol{H}$ and $\boldsymbol{A}$ and two radar vegetation-related features, i.e. pedestal height ($\boldsymbol{\psi}$) and radar vegetation index (RVI) (Tan, Ewe, and Chuah 2011; Tamiminia et al. 2017).

**Table 4.** Radar features extracted from two UAVSAR images.

| Polarimetric features and definitions |
|---|
| $\sigma_{HH} = 10\ln_{10}\lvert S_{HH}\rvert^2, \sigma_{HV} = 10\ln_{10}\lvert S_{HV}\rvert^2, \sigma_{VV} = 10\ln_{10}\lvert S_{VV}\rvert^2,$ |
| $\sigma_{RR} = 10\ln_{10}\lvert S_{RR}\rvert^2, \sigma_{RL} = 10\ln_{10}\lvert S_{RL}\rvert^2, \sigma_{LL} = 10\ln_{10}\lvert S_{LL}\rvert^2$ |
| $\{H : \text{horizontal},\ V : \text{vertical},\ R : \text{right} - \text{handed},\ L : \text{left} - \text{handed}\}$ |
| $R_{HHVV} = 10\ln_{10}\left(\frac{\lvert S_{HH}\rvert^2}{\lvert S_{VV}\rvert^2}\right), R_{HVHH} = 10\ln_{10}\left(\frac{\lvert S_{HV}\rvert^2}{\lvert S_{HH}\rvert^2}\right), R_{HVVV} = 10\ln_{10}\left(\frac{\lvert S_{HV}\rvert^2}{\lvert S_{VV}\rvert^2}\right), R_{RRLL} = 10\ln_{10}\left(\frac{\lvert S_{RR}\rvert^2}{\lvert S_{LL}\rvert^2}\right),$ |
| $R_{RLRR} = 10\ln_{10}\left(\frac{\lvert S_{RL}\rvert^2}{\lvert S_{RR}\rvert^2}\right), R_{RLLL} = 10\ln_{10}\left(\frac{\lvert S_{RL}\rvert^2}{\lvert S_{LL}\rvert^2}\right)$ |
| $R_{HH} = 10\ln_{10}\left(\frac{\lvert S_{HH}\rvert^2}{\text{span}}\right), R_{HV} = 10\ln_{10}\left(\frac{\lvert S_{HV}\rvert^2}{\text{span}}\right), R_{VV} = 10\ln_{10}\left(\frac{\lvert S_{VV}\rvert^2}{\text{span}}\right), R_{RR} = 10\ln_{10}\left(\frac{\lvert S_{RR}\rvert^2}{\text{span}}\right),$ |
| $R_{RL} = 10\ln_{10}\left(\frac{\lvert R_{EL}\rvert^2}{\text{span}}\right), R_{LL} = 10\ln_{10}\left(\frac{\lvert S_{LL}\rvert^2}{\text{span}}\right) \quad \{\text{span} = \lvert S_{HH}\rvert^2 + 2\lvert S_{HV}\rvert^2 + \lvert S_{VV}\rvert^2 = \lvert S_{RR}\rvert^2 + 2\lvert S_{RL}\rvert^2 + \lvert S_{LL}\rvert^2\}$ |
| $\rho_{HHVV} = \left\lvert\frac{S_{HH}.S_{VV}^*}{\sqrt{(S_{HH}.S_{HH}^*)(S_{VV}.S_{VV}^*)}}\right\rvert, \rho_{HVHH} = \left\lvert\frac{S_{HV}.S_{HH}^*}{\sqrt{(S_{HV}.S_{HV}^*)(S_{HH}.S_{HH}^*)}}\right\rvert, \rho_{HVVV} = \left\lvert\frac{S_{HV}.S_{VV}^*}{\sqrt{(S_{HV}.S_{HV}^*)(S_{VV}.S_{VV}^*)}}\right\rvert,$ |
| $\rho_{RRLL} = \left\lvert\frac{S_{RR}.S_{LL}^*}{\sqrt{(S_{RR}.S_{RR}^*)(S_{LL}.S_{LL}^*)}}\right\rvert, \rho_{RLRR} = \left\lvert\frac{S_{RL}.S_{RR}^*}{\sqrt{(S_{RL}.S_{RL}^*)(S_{RR}.S_{RR}^*)}}\right\rvert, \rho_{RLLL} = \left\lvert\frac{S_{RL}.S_{LL}^*}{\sqrt{(S_{RL}.S_{RL}^*)(S_{LL}.S_{LL}^*)}}\right\rvert$ |
| $\{* : \text{complexconjugate}, . : \text{vectorscalarordotproduct}\}$ |
| $\lambda_1, \lambda_2, \lambda_3, H = -\sum_{i=1}^{3} p_i \ln_3\ p_i, A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3}, \bar{\alpha} = \sum_{i=1}^{3} p_i \alpha_i \left\{p_i = \lambda_i / \lambda_i \sum_{k=1}^{3} \lambda_k - \sum_{k=1}^{3} \lambda_k\right\}$ |
| $HA, H(1-A), (1-H)A, (1-H)(1-A), \psi = \frac{\min(\lambda_1, \lambda_2, \lambda_3)}{\lambda_1 + \lambda_2 + \lambda_3}, \text{RVI} = \frac{4\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$ |
| $\lvert\alpha\rvert^2 = \left\lvert\frac{S_{HH} + S_{VV}}{\sqrt{2}}\right\rvert^2, \lvert\beta\rvert^2 = \left\lvert\frac{S_{HH} - S_{VV}}{\sqrt{2}}\right\rvert^2, \lvert\gamma\rvert^2 = 2\lvert S_{HV}\rvert^2, \lvert k_s\rvert^2 = \lvert S_{RL}\rvert^2, \lvert k_d\rvert^2 = \min\left(\lvert S_{RR}\rvert^2, \lvert S_{LL}\rvert^2\right),$ |
| $\lvert k_h\rvert^2 = \text{abs}\left(\lvert S_{RR}\rvert^2 - \lvert S_{LL}\rvert^2\right)$ |
| $\{s : \text{surface scattering},\ d : \text{double} - \text{bounce scattering},\ h : \text{helix scattering}\}$ |
| $P_s = f_s(1 + \lvert\beta\rvert^2), P_d = f_d(1 + \lvert\alpha\rvert^2), P_v = f_v, P_s^Y = f_s(1 + \lvert\beta\rvert^2), P_d^Y = f_d(1 + \lvert\alpha\rvert^2), P_v^Y = f_v, P_c^Y = f_c$ |
| $\{v : \text{volume scattering},\ c : \text{helix scattering}\}$ |

In addition to eigenvalue decompositions, coherent and incoherent target decompositions, which can offer useful information about the physics of scattering mechanisms, were extracted from the PolSAR imagery. We extracted three Pauli parameters ($|\boldsymbol{\alpha}|^2$, $|\boldsymbol{\beta}|^2$, and $|\boldsymbol{\gamma}|^2$) and three Krogager parameters ($|\boldsymbol{k}_s|^2$, $|\boldsymbol{k}_d|^2$, and $|\boldsymbol{k}_h|^2$) as coherent decompositions (Cloude and Pottier 1996; Krogager 1990). In addition, three Freeman-Durden parameters ($\boldsymbol{P}_s$, $\boldsymbol{P}_d$, and $\boldsymbol{P}_v$) and four Yamaguchi parameters ($\boldsymbol{P}_s{}^Y$, $\boldsymbol{P}_d{}^Y$, $\boldsymbol{P}_v{}^Y$, and $\boldsymbol{P}_c{}^Y$) were extracted as incoherent decompositions (Freeman and Durden 1998; Yamaguchi et al. 2005). Here, the subscripts 's', 'd', and 'v' denote the surface, double-bounce and volume scattering powers, respectively. Moreover, the subscripts 'h' and 'c' denote the helix scattering power for the Krogager and Yamaguchi (Y) decompositions, respectively. Table 4 provides a list of the 49 radar features mentioned earlier.

## 3. Methodology

### 3.1. *MSMD feature selection*

Most feature selection algorithms usually aim to choose the most relevant features with respect to the classes and to remove the redundant features (Zeng et al. 2015). The proposed MSMD method also has this goal, in order to reduce the dimensions of the data and to improve the classification accuracy. Figure 2 illustrates the two main steps of the MSMD method. These two steps are: selecting radar features which have the highest separability values (distance-based measure) and selecting optical features which have the lowest dependency upon each other (correlation-based measure). Accordingly, the MSMD method is both a distance-based and a correlation-based feature selection method.

#### 3.1.1. *Selecting the most relevant features*

The first step is to choose the most relevant features. For this purpose, we use the concept of the separability of features among the classes. Radar Earth observations have been shown to be very promising for analysing vegetation compared to optical observations (Joshi et al. 2016). Therefore, the separability values between pairwise classes for each radar feature are determined by an effective measure. In several recent studies, the Jeffries-Matusita (JM) distance has been utilized as a useful measure of separation in the classification task (Poth et al. 2001; Sousa, Pereira, and Silva 2003; Ifarraguerri and Prairie 2004; Li et al. 2013; Zhang, Feng, and Jiang 2010; Dabboor et al. 2014; Zhang et al. 2016; Lee, Kastens, and Egbert 2016; Wei et al. 2017). JM distance determines the separability of two classes for a feature as follows:

$$\mathrm{JM}_j = 2(1 - \mathrm{e}^{-B_j}), \qquad B_j = \frac{1}{8}(\mu_{1j} - \mu_{2j})^2 \frac{2}{\sigma_{1j}^2 + \sigma_{2j}^2} + \frac{1}{2}\ln\left(\frac{\sigma_{1j}^2 + \sigma_{2j}^2}{2\sigma_{1j}\sigma_{2j}}\right), \tag{1}$$

where $B$ is the Bhattacharyya distance and $\mu_{ij}$ and $\sigma_{ij}$ are the mean and variance values of two given classes ($i = 1$ or $i = 2$ denotes the class) for feature $j$. Unlike $B$, JM has a finite dynamic range [0, 2]. These values allow a better comparison of feature analysis results for identifying which feature has the best separability (Nussbaum and Menz 2008). A value of zero indicates weak separation and a value of two indicates the highest separation (Dabboor et al. 2014).

For each radar feature, we computed the JM measure between the class pairs as follows: (1) wheat versus other classes, (2) canola versus others, (3) soybeans versus others, (4) oats versus others, (5) corn versus others, (6) peas versus others and finally (7) broadleaf versus others. After computing JM values for all pairwise classes, the average JM value over the classes is calculated for each radar feature. Finally, the radar features ($\boldsymbol{X}$) with an average JM value higher than a specified threshold (thr1) are selected:

$$X : \mathrm{JM} > \mathrm{thr1}. \tag{2}$$

#### 3.1.2. *Removing redundant features*

After selecting the most relevant radar features, the second step is to eliminate the redundant features. For this purpose, we remove the optical features which have high

dependency with respect to the radar features selected in step 1, and select the remaining optical features. Thus, the correlation coefficient (CC) values among the optical and the selected radar features are computed as (Soja, Persson, and Ulander 2015):

$$\mathrm{CC} = \frac{E(\boldsymbol{XY})}{\sqrt{E(\boldsymbol{X^2}) \times E(\boldsymbol{Y^2})}}, \tag{3}$$

where $E()$ is the expectation operator and $\boldsymbol{X}$ and $\boldsymbol{Y}$ are radar and optical feature vectors respectively. CC has a range of [−1, 1]. A value close to zero indicates a low correlation and a value close to ±1 indicates a high correlation. We should select the optical features, which have low dependency with respect to the radar features. Thus, the optical features ($\boldsymbol{Y}$), which have an absolute CC value lower than a specified threshold (thr2) are selected in step 2:

$$\boldsymbol{Y}: \mathrm{abs}(\mathrm{CC}) < \mathrm{thr2}. \tag{4}$$

It is worth noting that an optical feature, which has a low correlation with one selected radar feature may have a high correlation with another selected radar feature. In this case, this optical feature should be removed from the selected feature subset.

In these two steps, we have chosen a set of radar and optical features which have the highest relevancy with respect to the classes and have the lowest redundancy among themselves.

### 3.2. *Random forest classifier*

The random forest (RF) classifier, as a state-of-the-art ensemble machine learning algorithm, was used in this study for crop mapping. RF is built based on a multitude of tree classifiers. The procedure is as follows. First, several training subsets are randomly sampled from the original training data. This method is typically called bootstrapping. For each training subset, several features are chosen from the entire set of features using a random subset feature selection method. Then, a decision-tree classifier is trained using each subset. In order to predict the label for a test sample, a class which is chosen by most of the decision tree classifiers is assigned (Breiman 2001; Pal 2005).

## 4. Implementation and results

The proposed MSMD method was compared with some well-known filter feature selection algorithms and a wrapper algorithm, with respect to accuracy and running time. For accuracy, we used the overall accuracy (OA) and the classification kappa coefficient. OA is a general metric for the evaluation of the classification and is the ratio of the diagonal pixels to all the samples of the confusion matrix. The kappa coefficient is computed to determine whether the values in a confusion matrix are significantly better than the values in a random assignment (Congalton 1991).

Filter methods utilize correlation, consistency, information and distance measures, and wrapper methods use a classifier to evaluate the selected feature subsets (Zeng

**Table 5.** OA values and kappa coefficients obtained by feature selection algorithms.

| | Method | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | MSMD | CFS | FCBF | SD | Fisher | Relief | JMIM | MRMR | GA-SVM |
| OA | **90.11** | 87.24 | 86.07 | 81.95 | 83.11 | 80.43 | 86.79 | 86.93 | 89.89 |
| Kappa coefficient | **0.88** | 0.84 | 0.83 | 0.79 | 0.80 | 0.77 | 0.84 | 0.84 | 0.88 |

et al. 2015). The filter methods used were: (1) correlation-based algorithms, e.g. correlation-based feature selection (CFS), fast correlation-based filter (FCBF) and statistical dependency (SD), (2) distance-based algorithms, e.g. Fisher and Relief algorithms, and (3) information-based algorithms, e.g. joint mutual information maximization (JMIM), and minimum redundancy maximum relevancy (MRMR) algorithms. The wrapper method utilized in this article was the genetic algorithm and Support Vector Machine (GA-SVM) method.

Table 5 shows that the proposed MSMD algorithm could produce higher OA values and kappa coefficients than all the filter algorithms and the wrapper algorithm. Compared to the correlation-based filter algorithms, i.e. CFS, FCBF, and SD, the MSMD method obtained a 3–10% higher OA. Moreover, the success of MSMD compared to the distance-based filter algorithms, i.e. the Fisher and Relief algorithms, was even more marked, with a 7–10% higher OA. The OA of MSMD was also 3% higher than that of the information-based filter algorithms, i.e. MRMR and JMIM.

Furthermore, the proposed MSMD algorithm obtained an OA and a kappa coefficient fairly close to those of the wrapper feature selection algorithm, i.e. GA-SVM. In the literature, wrapper feature selection algorithms produce higher accuracies than filter algorithms due to the direct use of a classifier (Zeng et al. 2015). In several studies, the accuracies obtained by the GA-SVM feature selection method were higher compared to other methods such as GA-K-nearest neighbour (GA-KNN), GA-multilayer perceptron (GA-MLP), MLP, SVM, KNN, etc. (e.g. Pal 2006; Bazi and Melgani 2006; Ghoggali, Melgani, and Bazi 2009; Tan et al. 2009; Fernandez-Lozano et al. 2013; De Stefano et al. 2014; Fassnacht et al. 2014; Gunavathi and Premalatha 2014; Salehi, Sahebi, and Maghsoudi 2014; Ghamisi and Benediktsson 2015; Sukawattanavijit, Chen, and Zhang 2017).

Figure 3 shows that the proposed MSMD method operated almost as fast as all the filter feature selection algorithms, and it was approximately 10 times faster than the wrapper algorithm. The higher speed and comparable accuracy of MSMD compared to the wrapper algorithm represents an enormous success for the proposed feature selection method. Furthermore, the proposed method had a much simpler structure compared to the GA-SVM method for feature selection.

Regarding visual comparisons, the classification map obtained by the MSMD method had a more homogeneous appearance than all the maps of the other feature selection methods, even that of the GA-SVM method. Figure 4 shows that corn, canola, wheat and peas were detected more perfectly and comprehensively in the MSMD map. Moreover, a mixture of broadleaf and other classes existed in some maps, such as those of the Fisher, SD, Relief and GA-SVM methods, while this problem was not observed in the MSMD map. Additionally, mixed pixels of oats and wheat, which were commonly seen in the maps of most of the other methods, rarely existed in the MSMD map. In the classified maps of methods such as FCBF, MRMR, Fisher, SD and Relief, a large number of soybeans

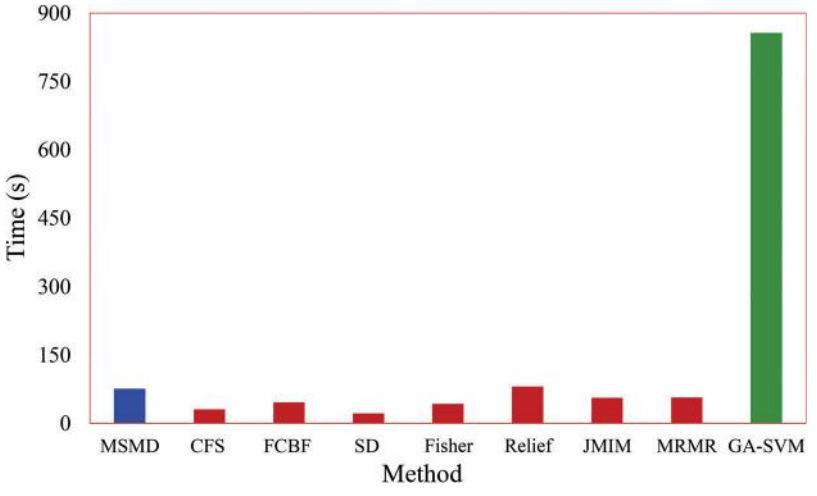


**Figure 3.** The running time of feature selection methods (red bars: filter methods, green bar: wrapper method, and blue bar: proposed method).

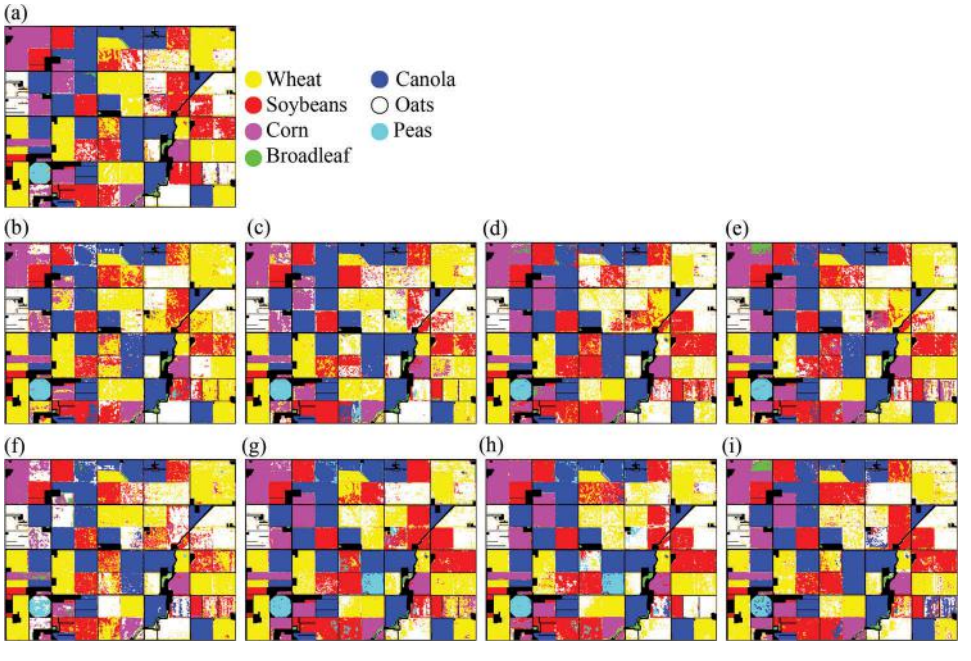


**Figure 4.** Classification maps obtained by: (a) MSMD, (b) CFS, (c) FCBF, (d) SD, (e) Fisher, (f) Relief, (g) JMIM, (h) MRMR, and (i) GA-SVM.

pixels, the equivalent of a farm, were wrongly classified as oats or other classes. Moreover, in the maps of information-based filter methods, i.e. MRMR and JMIM, a large number of canola pixels, the equivalent of a farm, were wrongly identified as peas, but both these two problems were resolved in the MSMD map.

It should be noted that the performance of the MSMD method depends on the values of the two thresholds, especially thr1. A sensitivity analysis was performed to specify the

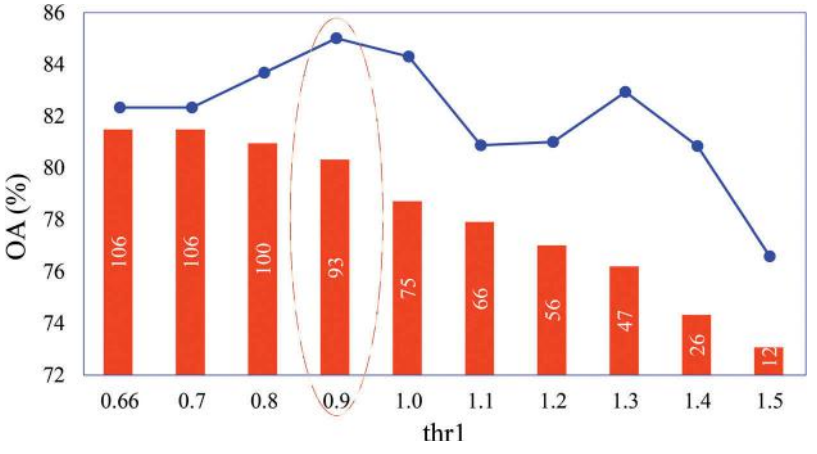


**Figure 5.** The impact of thr1 values on OA of MSMD method. The blue line shows the OA value and the red bars show the number of selected features for each thr1 value. The circle shows the optimum value for thr1 due to higher OA.

impact of the thr1 value on the OA and the number of selected features. Figure 5 indicates that the number of selected features decreased when thr1 increased. However, the OA values increased with values of thr1 up to 0.9, and then decreased (except for thr1 = 1.3). Therefore, we selected the optimum value of 0.9 for thr1. Using this threshold, 93 features, i.e. almost half the features, were selected.

Figure 5 also shows that the differences in OA between this optimum value and other values (for thr1 < 1.5) are in the range 1–5%. This fact implies that an imprecise selection of the thr1 value will cost up to 5% of the overall accuracy. Furthermore, we found that thr1 = 0.7, i.e. 0.2 less than the optimum value, may be specified by Otsu's method from JM values.

Otsu's method assumes that the data containes two classes. It then calculates the optimum threshold separating the two classes so that their intra-class variance is minimal, or equivalently so that their inter-class variance is maximal (Otsu 1979). The OA obtained using this threshold was almost 3% lower than the OA for the optimum value. Thus, we can conclude that Otsu's method of threshold calculation can lead to a value close to the optimum thr1 value. Figure 5 also indicated that when thr1 was 1.5, the OA could be 8.5% lower than the optimum value. As a result, a much higher value of thr1 may decrease the accuracy and degrade the results.

Another experimental test performed in this study was to determine the optimum features for each class using the MSMD method, by setting a particular threshold for the separability value for that class. Table 6 presents the *F*-score values of seven classes obtained from the optimum features for each class. The *F*-score is the harmonic mean of the producer's and user's accuracies for each class (Powers 2011).

From Table 6, it can be seen that the canola and peas classes obtained the highest *F*-score values (>95%) with the smallest number of radar and optical features (30). By contrast, oats obtained the lowest *F*-score value (≈75%) with a higher number of features compared to canola and peas. Similar optimum features selected for all the classes are presented in Table 7.

**Table 6.** *F*-score values obtained for each class by the best features.

| | Class | | | | | | |
|---|---|---|---|---|---|---|---|
| | Wh | Ca | So | Oa | Co | Pe | Br |
| *F*-score | 85.06 | 98.06 | 92.64 | 75.10 | 92.94 | 95.45 | 82.43 |
| Number of selected features | 68 | 30 | 62 | 62 | 48 | 30 | 85 |

**Table 7.** Similar features selected for all the classes.

| Features | Selected features | Number |
|---|---|---|
| Radar features | $\sigma^1_{vv}$, $\sigma^1_{rr}$, $\sigma^1_{ll}$, $\lvert\beta\rvert^{2^1}$, $\lvert\gamma\rvert^{2^1}$, $\lvert k_s\rvert^{2^1}$, $\lvert k_d\rvert^{2^1}$ | 7 |
| | $\sigma^2_{hh}$, $\sigma^2_{hv}$, $\sigma^2_{rl}$, $\lvert\alpha\rvert^{2^2}$, $\lvert\gamma\rvert^{2^2}$ | 5 |
| Optical features | $SAVI^1$, $HOM^1_{PC1}$ | 2 |
| | $\sigma^2_{PC2}$, $CON^2_{PC2}$ | 2 |
| Superscript 1: date1; superscript 2: date2 | | 16 |

Table 7 confirms that among the radar features on the first date, the volume scattering components played the chief role in the successful classification of all the crop types. By contrast, on the second date, the odd-bounce and even-bounce features had the most important roles. Furthermore, this table indicates that the SAVI was the best VI used in this study. Among the textural indicators, homogeneity, variance, and contrast had the most significant roles in improving the detection of all the crop types.

A further conclusion was that the oats and wheat had a significant number of similar optimum features. This is logical and acceptable since these two crop types usually have the same reflectance and structural behaviour. However, wheat has a 10% higher detection rate than oats. Moreover, we can resolve the mixture of oats and wheat using the different optimum features of each crop. This is true for all crop types.

Another point to be noted here is that we also designed the MSMD algorithm in reverse. We employed the first step of MSMD on optical images and then the second step on radar images. The results were not as good as those of the current MSMD, and the OA was 5% lower. This fact confirmed that radar data are more relevant for analysing vegetation than optical data, and the separability analysis should be performed for radar features.

## 5. Conclusion

In this research, we proposed an intrinsic feature selection method from a temporal set of radar and optical features for crop mapping. First, the radar features which had the highest separability among the classes were selected. Then, the optical features which had the lowest correlation with the radar features were selected. The proposed method was compared with several filter and wrapper feature selection algorithms. The results of the experimental tests indicated that the MSMD method had a higher accuracy and an equal running time compared to the filter methods. Furthermore, it had a slightly better accuracy and a much lower running time than the wrapper algorithm. Another advantage of this approach over the wrapper algorithm is its simpler structure for selecting the optimum features.

The main advantage of the MSMD method compared to the other feature selection methods is its ability to choose the optimum features for each class by determining a

specific threshold for the separability value for that class. In this way, we can find similar or different optimum features for the various classes and can resolve their confusion by weighting the features.

This idea of extracting the best features from the freely available Earth observations from Sentinel-1 and -2, Landsat 8, and the future RADARSAT Constellation Mission, seems to be very helpful for operational crop-mapping activities.

## Acknowledgements

The authors would like to present their acknowledgments to the JPL NASA, the SMAPVEX 2012 team, and the Agriculture and Agri-Food for providing for the PolSAR and the optical data used in this research.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

Alberga, V. 2007. "A Study of Land Cover Classification Using Polarimetric SAR Parameters." *International Journal of Remote Sensing* 28 (17): 3851–3870. doi:10.1080/01431160601075541.

Alberga, V., G. Satalino, and D. K. Staykova. 2008. "Comparison of Polarimetric SAR Observables in Terms of Classification Performance." *International Journal of Remote Sensing* 29 (14): 4129–4150. doi:10.1080/01431160701840182.

Bazi, Y., and F. Melgani. 2006. "Toward an Optimal SVM Classification System for Hyperspectral Remote Sensing Images." *IEEE Transactions on Geoscience and Remote Sensing* 44 (11): 3374–3385. doi:10.1109/TGRS.2006.880628.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/A:1010933404324.

Chen, S. W., Y. Z. Li, and X. S. Wang. 2015. "Crop Discrimination Based on Polarimetric Correlation Coefficients Optimization for PolSAR Data." *International Journal of Remote Sensing* 36 (16): 4233–4249. doi:10.1080/01431161.2015.1079345.

Cloude, S., and E. Pottier. 1996. "A Review of Target Decomposition Theorems in Radar Polarimetry." *IEEE Transactions on Geoscience and Remote Sensing* 34 (2): 498–518. doi:10.1109/36.485127.

Congalton, R. G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034-4257(91)90048-B.

Dabboor, M., S. Howell, M. Shokr, and J. Yackel. 2014. "The Jeffries–Matusita Distance for the Case of Complex Wishart Distribution as a Separability Criterion for Fully Polarimetric SAR Data." *International Journal of Remote Sensing* 35 (19): 6859–6873.

De Souza, C. H. W., E. Mercante, J. A. Johann, R. A. C. Lamparelli, and M. A. Uribe-Opazo. 2015. "Mapping and Discrimination of Soya Bean and Corn Crops Using Spectro-Temporal Profiles of Vegetation Indices." *International Journal of Remote Sensing* 36 (7): 1809–1824. doi:10.1080/01431161.2015.1026956.

De Stefano, C., F. Fontanella, C. Marrocco, and A. S. Di Freca. 2014. "A GA-Based Feature Selection Approach with an Application to Handwritten Character Recognition." *Pattern Recognition Letters* 35: 130–141. doi:10.1016/j.patrec.2013.01.026.

Entezari, I., M. Motagh, and B. Mansouri. 2012. "Comparison of the Performance of L-Band Polarimetric Parameters for Land Cover Classification." *Canadian Journal of Remote Sensing* 38 (5): 629–643. doi:10.5589/m12-051.

Fassnacht, F. E., C. Neumann, M. Förster, H. Buddenbaum, A. Ghosh, A. Clasen, P. K. Joshi, and B. Koch. 2014. "Comparison of Feature Reduction Algorithms for Classifying Tree Species with

Hyperspectral Data on Three Central European Test Sites." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (6): 2547–2561. doi:10.1109/JSTARS.2014.2329390.

Fernandez-Lozano, C., C. Canto, M. Gestal, J. M. Andrade-Garda, J. R. Rabuñal, J. Dorado, and A. Pazos. 2013. "Hybrid Model Based on Genetic Algorithms and SVM Applied to Variable Selection within Fruit Juice Classification." *The Scientific World Journal* 2013: 1–13. doi:10.1155/2013/982438.

Foody, G. M., M. B. Mcculloch, and W. B. Yates. 1994. "Crop Classification from C-Band Polarimetric Radar Data." *International Journal of Remote Sensing* 15 (14): 2871–2885. doi:10.1080/01431169408954289.

Forkuor, G., C. Conrad, M. Thiel, T. Ullmann, and E. Zoungrana. 2014. "Integration of Optical and Synthetic Aperture Radar Imagery for Improving Crop Mapping in Northwestern Benin, West Africa." *Remote Sensing* 6 (7): 6472–6499. doi:10.3390/rs6076472.

Freeman, A., and S. L. Durden. 1998. "A Three-Component Scattering Model for Polarimetric SAR Data." *IEEE Transactions on Geoscience and Remote Sensing* 36 (3): 963–973. doi:10.1109/36.673687.

Ghamisi, P., and J. A. Benediktsson. 2015. "Feature Selection Based on Hybridization of Genetic Algorithm and Particle Swarm Optimization." *IEEE Geoscience and Remote Sensing Letters* 12 (2): 309–313. doi:10.1109/LGRS.2014.2337320.

Ghoggali, N., F. Melgani, and Y. Bazi. 2009. "A Multiobjective Genetic SVM Approach for Classification Problems with Limited Training Samples." *IEEE Transactions on Geoscience and Remote Sensing* 47 (6): 1707–1718. doi:10.1109/TGRS.2008.2007128.

Gunavathi, C., and K. Premalatha. 2014. "Performance Analysis of Genetic Algorithm with KNN and SVM for Feature Selection in Tumor Classification." *International Journal of Computer, Electrical, Automation, Control and Information Engineering* 8 (8): 1490–1497.

Haldar, D., and C. Patnaik. 2010. "Synergistic Use of Multi-Temporal Radarsat SAR and AWiFS Data for Rabi Rice Identification." *Journal of the Indian Society of Remote Sensing* 38 (1): 153–160. doi:10.1007/s12524-010-0006-x.

Hoang, H. K., M. Bernier, S. Duchesne, and Y. M. Tran. 2016. "Rice Mapping Using RADARSAT-2 Dual-And Quad-Pol Data in a Complex Land-Use Watershed: Cau River Basin (Vietnam)." *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 9 (7): 3082–3096. doi:10.1109/JSTARS.2016.2586102.

Hosseini, M., H. McNairn, A. Merzouki, and A. Pacheco. 2015. "Estimation of Leaf Area Index (LAI) in Corn and Soybeans Using Multi-Polarization C- and L-Band Radar Data." *Remote Sensing of Environment* 170: 77–89. doi:10.1016/j.rse.2015.09.002.

Ifarraguerri, A., and M. W. Prairie. 2004. "Visual Method for Spectral Band Selection." *IEEE Geoscience and Remote Sensing Letters* 1 (2): 101–106. doi:10.1109/LGRS.2003.822879.

Inglada, J., A. Vincent, M. Arias, and C. Marais-Sicre. 2016. "Improved Early Crop Type Identification by Joint Use of High Temporal Resolution SAR and Optical Image Time Series." *Remote Sensing* 8 (5): 362. doi:10.3390/rs8050362.

Joshi, N., M. Baumann, A. Ehammer, R. Fensholt, K. Grogan, P. Hostert, M. R. Jepsen, et al. 2016. "A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring." *Remote Sensing* 8 (1): 70, 1–23. doi:10.3390/rs8010070.

Khosravi, I., A. Safari, S. Homayouni, and H. McNairn. 2017. "Enhanced Decision Tree Ensembles for Land Cover Mapping from Fully Polarimetric SAR Data." *International Journal of Remote Sensing* 38 (23): 7138–7160. doi:10.1080/01431161.2017.1372863.

Kim, H. O., and J. M. Yeom. 2014. "Effect of Red-Edge and Texture Features for Object-Based Paddy Rice Crop Classification Using RapidEye Multi-Spectral Satellite Image Data." *International Journal of Remote Sensing* 35 (19): 7046–7068.

Krogager, E. 1990. "A New Decomposition of the Radar Target Scattering Matrix." *Electronic Letters* 26 (18): 1525–1527. doi:10.1049/el:19900979.

Kross, A., H. McNairn, D. Lapen, M. Sunohara, and C. Champagne. 2015. "Assessment of RapidEye Vegetation Indices for Estimation of Leaf Area Index and Biomass in Corn and Soybean Crops."

*International Journal of Applied Earth Observation and Geoinformation* 34: 235–248. doi:10.1016/j.jag.2014.08.002.

Kussul, N., G. Lemoine, F. J. Gallego, S. V. Skakun, M. Lavreniuk, and A. Y. Shelestov. 2016. "Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data." *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 9 (6): 2500–2508. doi:10.1109/JSTARS.2016.2560141.

Kussul, N., S. Skakun, O. Kravchenko, A. Shelestov, J. F. Gallego, and O. Kussul. 2013. "Application of Satellite Optical and SAR Images for Crop Mapping and Area Estimation in Ukraine." *International Journal of Information Modelling Analysis* 7: 203–211.

Lardeux, C., P.-L. Frison, C. Tison, J.-C. Souyris, B. Stoll, B. Fruneau, and J.-P. Rudant. 2009. "Support Vector Machine for Multifrequency SAR Polarimetric Data Classification." *IEEE Transactions on Geoscience and Remote Sensing* 47 (12): 4143–4152. doi:10.1109/TGRS.2009.2023908.

Lee, E., J. H. Kastens, and S. L. Egbert. 2016. "Investigating Collection 4 versus Collection 5 MODIS 250 M NDVI Time-Series Data for Crop Separability in Kansas, USA." *International Journal of Remote Sensing* 37 (2): 341–355. doi:10.1080/01431161.2015.1125556.

Li, X., Q. Meng, X. Gu, T. Jancso, T. Yu, K. Wang, and S. Mavromatis. 2013. "A Hybrid Method Combining Pixel-Based and Object-Oriented Methods and Its Application in Hungary Using Chinese HJ-1 Satellite Images." *International Journal of Remote Sensing* 34 (13): 4655–4668. doi:10.1080/01431161.2013.780669.

Lönnqvist, A., Y. Rauste, M. Molinier, and T. Häme. 2010. "Polarimetric SAR Data in Land Cover Mapping in Boreal Zone." *IEEE Transactions on Geoscience and Remote Sensing* 48 (10): 3652–3662. doi:10.1109/TGRS.2010.2048115.

Löw, F., G. Schorcht, U. Michel, S. Dech, and C. Conrad. 2012. "Per-Field Crop Classification in Irrigated Agricultural Regions in Middle Asia Using Random Forest and Support Vector Machine Ensemble." In *Proceeding SPIE Earth Resources and Environmental Remote Sensing/GIS Applications III 85380R*. 8538: 1–11. doi: 10.1117/12.974588

Mahdian, M., S. Homayouni, M. A. Fazel, and F. Mohammadimanesh. 2013. "Agricultural Land Classification Based on Statistical Analysis of Full Polarimetric SAR Data." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL1/W3: 257–261. doi:10.5194/isprsarchives-XL-1-W3.

McNairn, H., C. Champagne, J. Shang, D. Holmstrom, and G. Reichert. 2009. "Integration of Optical and Synthetic Aperture Radar (SAR) Imagery for Delivering Operational Annual Crop Inventories." *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (5): 434–449. doi:10.1016/j.isprsjprs.2008.07.006.

McNairn, H., T. J. Jackson, G. Wiseman, S. Belair, A. Berg, P. Bullock, A. Colliander, et al. 2015. "The Soil Moisture Active Passive Validation Experiment 2012 (SMAPVEX12): Prelaunch Calibration and Validation of the SMAP Soil Moisture Algorithms." *IEEE Transactions on Geoscience and Remote Sensing* 53 (5): 2784–2801. doi:10.1109/TGRS.2014.2364913.

Nussbaum, S., and G. Menz. 2008. "Object-based Image Analysis and Treaty Verification: New Approaches in Remote Sensing-applied to Nuclear Facilities in Iran." *Springer Science & Business Media* pp. 51–55. doi: 10.1007/978-1-4020-6961-1.

Otsu, N. 1979. "A Threshold Selection Method from Gray-Level Histograms." *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1): 62–66. doi:10.1109/TSMC.1979.4310076.

Pal, M. 2005. "Random Forest Classifier for Remote Sensing Classification." *International Journal of Remote Sensing* 26 (1): 217–222. doi:10.1080/01431160412331269698.

Pal, M. 2006. "Support Vector Machine-Based Feature Selection for Land Cover Classification: A Case Study with DAIS Hyperspectral Data." *International Journal of Remote Sensing* 27 (14): 2877–2894. doi:10.1080/01431160500242515.

Park, S., and J. Im. 2016. "Classification of Croplands through Fusion of Optical and SAR Time Series Data." *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B7: 703–704. doi:10.5194/isprsarchives-XLI-B7-703-2016.

Peña-Barragán, J. M., M. K. Ngugi, R. E. Plant, and J. Six. 2016. "Object-Based Crop Identification Using Multiple Vegetation Indices, Textural Features and Crop Phenology." *Remote Sensing of Environment* 115 (6): 1301–1316. doi:10.1016/j.rse.2011.01.009.

Poth, A., D. Klaus, M. Voss, and G. Stein. 2001. "Optimization at Multi-Spectral Land Cover Classification with Fuzzy Clustering and the Kohonen Feature Map." *International Journal of Remote Sensing* 22 (8): 1423–1439. doi:10.1080/01431160117710.

Powers, D. M. 2011. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation." *Journal of Machine Learning Technologies* 2 (1): 37–63.

Rao, P. N., M. S. Sai, K. Sreenivas, M. K. Rao, B. R. M. Rao, R. S. Dwivedi, and L. Venkataratnam. 2002. "Textural Analysis of IRS-1D Panchromatic Data for Land Cover Classification." *International Journal of Remote Sensing* 23 (17): 3327–3345. doi:10.1080/01431160110104665.

Salehi, M., M. R. Sahebi, and Y. Maghsoudi. 2014. "Improving the Accuracy of Urban Land Cover Classification Using Radarsat-2 PolSAR Data." *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 7 (4): 1394–1401. doi:10.1109/JSTARS.2013.2273074.

Siachalou, S., G. Mallinis, and M. Tsakiri-Strati. 2015. "A Hidden Markov Models Approach for Crop Classification: Linking Crop Phenology to Time Series of Multi-Sensor Remote Sensing Data." *Remote Sensing* 7 (4): 3633–3650. doi:10.3390/rs70403633.

Skakun, S., N. Kussul, A. Y. Shelestov, M. Lavreniuk, and O. Kussul. 2016. "Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine." *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 9 (8): 3712–3719. doi:10.1109/JSTARS.2015.2454297.

Soja, M. J., H. Persson, and L. M. Ulander. 2015. "Estimation of Forest Height and Canopy Density from a Single InSAR Correlation Coefficient." *IEEE Geoscience and Remote Sensing Letters* 12 (3): 646–650. doi:10.1109/LGRS.2014.2354551.

Soria-Ruiz, J., Y. Fernandez-Ordonez, and I. H. Woodhouse. 2010. "Land-Cover Classification Using Radar and Optical Images: A Case Study in Central Mexico." *International Journal of Remote Sensing* 31 (12): 3291–3305. doi:10.1080/01431160903160777.

Sousa, A. M. O., J. M. C. Pereira, and J. M. N. Silva. 2003. "Evaluating the Performance of Multitemporal Image Compositing Algorithms for Burned Area Analysis." *International Journal of Remote Sensing* 24 (6): 1219–1236. doi:10.1080/01431160110114466.

Sukawattanavijit, C., J. Chen, and H. Zhang. 2017. "GA-SVM Algorithm for Improving Land-Cover Classification Using SAR and Optical Remote Sensing Data." *IEEE Geoscience and Remote Sensing Letters* 14 (3): 284–288. doi:10.1109/LGRS.2016.2628406.

Tamiminia, H., S. Homayouni, H. McNairn, and A. Safari. 2017. "A Particle Swarm Optimized Kernel-Based Clustering Method for Crop Mapping from Multi-Temporal Polarimetric L-Band SAR Observations." *International Journal of Applied Earth Observation and Geoinformation* 58: 201–212. doi:10.1016/j.jag.2017.02.010.

Tan, C. P., H. T. Ewe, and H. T. Chuah. 2011. "Agricultural Crop-Type Classification of Multi-Polarization SAR Images Using a Hybrid Entropy Decomposition and Support Vector Machine Technique." *International Journal of Remote Sensing* 32 (22): 7057–7071. doi:10.1080/01431161.2011.613414.

Tan, K. C., E. J. Teoh, Q. Yu, and K. C. Goh. 2009. "A Hybrid Evolutionary Algorithm for Attribute Selection in Data Mining." *Expert Systems with Applications* 36 (4): 8616–8630. doi:10.1016/j.eswa.2008.10.013.

Theodoridis, S., and K. Koutroumbas. 2003. *Pattern Recognition*. 2nd ed. New York, NY: Academic Press.

Torbick, N., W. A. Salas, S. Hagen, and X. Xiao. 2011. "Monitoring Rice Agriculture in the Sacramento Valley, USA with Multitemporal PALSAR and MODIS Imagery." *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 4 (2): 451–457. doi:10.1109/JSTARS.2010.2091493.

Villa, P., D. Stroppiana, G. Fontanelli, R. Azar, and P. A. Brivio. 2015. "In-Season Mapping of Crop Type with Optical and X-Band SAR Data: A Classification Tree Approach Using Synoptic Seasonal Features." *Remote Sensing* 7 (10): 12859–12886. doi:10.3390/rs71012859.

Wardlow, B. D., and S. L. Egbert. 2010. "A Comparison of MODIS 250-M EVI and NDVI Data for Crop Mapping: A Case Study for Southwest Kansas." *International Journal of Remote Sensing* 31 (3): 805–830. doi:10.1080/01431160902897858.

Wei, Z., H. He, H. Hao, and W. Gao. 2017. "Automated Mapping of Landforms through the Application of Supervised Classification to LiDAR-Derived DEMs and the Identification of Earthquake Ruptures." *International Journal of Remote Sensing* 38 (23): 7196–7219. doi:10.1080/01431161.2017.1372861.

Whelen, T., and P. Siqueira. 2017. "Use of Time-Series L-Band UAVSAR Data for the Classification of Agricultural Fields in the San Joaquin Valley." *Remote Sensing of Environment* 193: 216–224. doi:10.1016/j.rse.2017.03.014.

Xie, H., Y. Q. Tian, J. A. Granillo, and G. R. Keller. 2007. "Suitable Remote Sensing Method and Data for Mapping and Measuring Active Crop Fields." *International Journal of Remote Sensing* 28 (2): 395–411. doi:10.1080/01431160600702673.

Yamaguchi, Y., T. Moriyama, M. Ishido, and H. Yamada. 2005. "Four-Component Scattering Model for Polarimetric SAR Image Decomposition." *IEEE Transactions on Geoscience and Remote Sensing* 43 (8): 1699–1706. doi:10.1109/TGRS.2005.852084.

Yekkehkhany, B., A. Safari, S. Homayouni, and M. Hasanlou. 2014b. "A Comparison Study of Different Kernel Functions for SVM-Based Classification of Multi-Temporal Polarimetry SAR Data." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40 (2): 281. doi:10.5194/isprsarchives-XL-2-W3-281-2014.

Yekkehkhany, B., S. Homayouni, H. McNairn, and A. Safari. 2014a. "Multi-Temporal Full Polarimetry L-Band SAR Data Classification for Agriculture Land Cover Mapping." In *Geoscience and Remote Sensing Symposium (IGARSS)*, 2014 IEEE International, Quebec City, QC, Canada, 13-18 July 2014. 2770–2773. doi:10.1109/IGARSS.2014.6947050

Zeng, Z., H. Zhang, R. Zhang, and C. Yin. 2015. "A Novel Feature Selection Method Considering Feature Interaction." *Pattern Recognition* 48 (8): 2656–2666. doi:10.1016/j.patcog.2015.02.025.

Zhang, X., X. Feng, and H. Jiang. 2010. "Object-Oriented Method for Urban Vegetation Mapping Using IKONOS Imagery." *International Journal of Remote Sensing* 31 (1): 177–196. doi:10.1080/01431160902882603.

Zhang, X., Y. Sun, K. Shang, L. Zhang, and S. Wang. 2016. "Crop Classification Based on Feature Band Set Construction and Object-Oriented Approach Using Hyperspectral Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (9): 4117–4128. doi:10.1109/JSTARS.2016.2577339.

Zhu, L., and R. Tateishi. 2006. "Fusion of Multisensor Multitemporal Satellite Data for Land Cover Mapping." *International Journal of Remote Sensing* 27 (5): 903–918. doi:10.1080/01431160310001398l8.