Dual-polarimetric descriptors from Sentinel-1 GRD SAR data for crop growth assessment

Narayanarao Bhogapurapu [1], Subhadip Dey [2], Avik Bhattacharya [2], Dipankar Mandal [2], Juan M Lopez Sanchez [2], Heather McNairn [2], Carlos López-Martínez [2], and Y S Rao [2]

[1]Indian Institute of Technology Bombay
[2]Affiliation not available

October 30, 2023

Abstract

Accurate and high-resolution spatio-temporal information about crop phenology obtained from Synthetic Aperture Radar (SAR) data is an essential component for crop management and yield estimation at a local scale. Crop growth monitoring studies seldom exploit complete polarimetric information contained in dual-pol GRD SAR data. In this study, we propose three polarimetric descriptors: the pseudo scattering-type parameter (ϑc), the pseudo scattering entropy parameter (Hc), and the co-pol purity parameter (mc) from dual-pol S1 GRD SAR data. We also introduce a novel unsupervised clustering framework using Hc and ϑc with six clustering zones to represent various scattering mechanisms. We implemented the proposed algorithm on the cloud-based Google Earth Engine (GEE) platform for Sentinel-1 SAR data. We have shown the sensitivity of these descriptors over a time series of data for wheat and canola crops at a test site in Canada. From the leaf development stage to the flowering stage for both crops, the pseudo scattering-type parameter ϑc changes by approximately 17°. Moreover, within the entire phenology window, both mc and Hc varies by about 0.6. The effectiveness of ϑc and Hc to cluster the phenological stages for the two crops is also evident from the clustering plot. During the leaf development stage, about 90 % of the sampling points were clustered into the low to medium entropy scattering zone for both the crops. Throughout the flowering stage, the entire cluster shifted into the high entropy vegetation scattering zone. Finally, during the ripening stage, the clusters of sample points were split between the high entropy vegetation scattering zone and the high entropy distributed scattering zone, with > 55 % of the sampling points in the high entropy distributed scattering zone. This innovative clustering framework will facilitate the operational use of S1 GRD SAR data for agricultural applications.

This article is submitted to ISPRS Journal of Photogrammetry and Remote Sensing

Dual-polarimetric descriptors from Sentinel-1 GRD SAR data for crop growth assessment

Narayanarao Bhogapurapu[a,*], Subhadip Dey[a], Avik Bhattacharya[a], Dipankar Mandal[a], Juan M. Lopez-Sanchez[b], Heather McNairn[c], Carlos López-Martínez[d], Y. S. Rao[a]

[a]*Microwave Remote Sensing Lab, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India*
[b]*University of Alicante, Alicante, Spain*
[c]*Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Canada*
[d]*Signal Theory and Communications Department (TSC), Universitat Politécnica de Catalunya (UPC), Barcelona, Spain*

Abstract

Accurate and high-resolution spatio-temporal information about crop phenology obtained from Synthetic Aperture Radar (SAR) data is an essential component for crop management and yield estimation at a local scale. Crop growth monitoring studies seldom exploit complete polarimetric information contained in dual-pol GRD SAR data. In this study, we propose three polarimetric descriptors: the pseudo scattering-type parameter (θ_c), the pseudo scattering entropy parameter (H_c), and the co-pol purity parameter (m_c) from dual-pol S1 GRD SAR data. We also introduce a novel unsupervised clustering framework using H_c and θ_c with six clustering zones to represent various scattering mechanisms. We implemented the proposed algorithm on the cloud-based Google Earth Engine (GEE) platform for Sentinel-1 SAR data. We have shown the sensitivity of these descriptors over a time series of

*Corresponding author: N. Bhogapurapu (narayanarao.bhogapurapu@gmail.com)

data for wheat and canola crops at a test site in Canada. From the leaf development stage to the flowering stage for both crops, the pseudo scattering-type parameter θ_c changes by approximately 17°. Moreover, within the entire phenology window, both m_c and H_c varies by about 0.6. The effectiveness of θ_c and H_c to cluster the phenological stages for the two crops is also evident from the clustering plot. During the leaf development stage, about 90 % of the sampling points were clustered into the low to medium entropy scattering zone for both the crops. Throughout the flowering stage, the entire cluster shifted into the high entropy vegetation scattering zone. Finally, during the ripening stage, the clusters of sample points were split between the high entropy vegetation scattering zone and the high entropy distributed scattering zone, with $> 55\,\%$ of the sampling points in the high entropy distributed scattering zone. This innovative clustering framework will facilitate the operational use of S1 GRD SAR data for agricultural applications.
Keywords: GRD SAR, Dual-pol, phenology, Unsupervised clustering, GEE, Sentinel-1

---

1. Introduction

Synthetic Aperture Radar (SAR) data have been extensively used for crop growth monitoring and classification, yield estimation, and phenological stages characterization. This is due to their high sensitivity towards the structure and dielectric properties of crop canopies (Ulaby, 1975; Ulaby and El-Rayes, 1987; Brisco et al., 1992; Ferrazzoli et al., 1992; McNairn and Brisco, 2004; Steele-Dunne et al., 2017). Because of its high spatial resolution and all-weather capabilities, SAR has proven to be a promising data source for continuously monitoring crops at field scales. The interaction of the SAR signal with crop canopies and the underlying soil varies with wavelength, polarization and angle of incidence (Ferrazzoli et al., 1992; Davidson et al., 2000). In general, during the early vegetative growth stage, the SAR backscatter signal is significantly affected by the underlying soil (Wiseman et al., 2014). The canopy structure and canopy moisture distribution are among major observable biophysical parameters that influence backscatter at each phenological stage. Further, the dense and complex geometry of the canopy leads to randomness in the scattering, which is more significant for fully developed crop canopies (Mascolo et al., 2016; Hariharan et al., 2018; Wang et al., 2019). The scattering becomes increasingly unpredictable during fruit development stages, leading to greater randomness in the SAR response (Jiao et al., 2014).

The availability of dual-pol SAR data acquired by the Sentinel-1 constellation provides diverse opportunities for many crop monitoring applications (ESA, 2017). Compared to full-pol mode, dual-pol modes have advantages in terms of larger swath widths and lower data volumes, but at the expense of reduced polarimetric information (Lee et al., 2001; Ainsworth et al., 2009). The Sentinel-1 (S1) SAR sensor in Interferometric Wide (IW) swath mode acquires data in dual-polarization, either in VV-VH or HH-HV.

Several researchers indicated the potential use of dual-pol backscatter intensities for crop type identification (Kussul et al., 2016; Nguyen et al., 2016; Bargiel, 2017; Van Tricht et al., 2018; Mandal et al., 2018; Whelen and Siqueira, 2018; Minasny et al., 2019; Arias et al., 2020), crop biophysical parameter estimation (Bousbih et al., 2017; Kumar et al., 2018; Mandal et al., 2020a), and phenology identification (Nelson et al., 2014; De Bernardis et al., 2015; Lasko et al., 2018; Singha et al., 2019; Fikriyah et al., 2019). Cloude (2007) proposed a clustering technique for dual-polarimetric (HH-HV or VV-VH) SAR data. An eigendecomposition of the 2×2 covariance matrix is performed to characterize scattering mechanisms from targets. The average scattering angle $\overline{\alpha}$ is obtained from the two orthogonal polarization states weighted by their corresponding pseudo probabilities obtained from the eigenvalues. The entropy H is obtained from the pseudo probabilities. Ainsworth et al. (2008) introduced a scattering-type parameter θ for dual-pol SLC data (HH-HV) utilizing the eigendecomposition technique. This parameter is presented as a measure between the cross- and co-pol backscatter ratio ($\sigma^{\circ}_{\mathrm{XY}}/\sigma^{\circ}_{\mathrm{XX}}$). It was stated that although the formulation is similar to Cloude α, the scattering information content is different. Utilizing θ and the scattering entropy (H) for dual-pol SAR data, an unsupervised clustering framework was proposed to identify different targets based on their scattering mechanisms. The unsupervised clustering plane was divided into eight different zones based on the scattering types.

Besides this, several vegetation descriptors such as the Radar Vegetation Index (RVI) for dual-pol (Trudel et al., 2012), Dual-Pol SAR Vegetation Index (DPSVI) (Periasamy, 2018), and Dual-pol Radar Vegetation Index (DpRVI) (Mandal et al., 2020b) have been developed for crop growth monitoring and biophysical parameter retrieval. However, similar descriptors are not directly available for dual-pol GRD SAR data products.

In particular, investigation often is limited to the direct use of backscatter intensities or their ratios for crop phenology identification and clustering. Vreugdenhil et al. (2018) studied the sensitivity of backscatter intensities and the cross-pol ratio (VH/VV) to crop biophysical parameters such as the Vegetation Water Content (VWC), Leaf Area Index (LAI), biomass, and the plant height for three different crops using the Sentinel-1 GRD SAR data.

Temporal sensitivity analysis using various machine learning models has shown that the cross-pol ratio is a valuable parameter for monitoring crop biophysical parameters and phenology. Song and Wang (2019) analyzed the temporal trend of VV and VH backscatter intensities to identify and map winter wheat crop using a parallelepiped classifier. They distinguished different phenology stages by exploring the temporal trend of the VH/VV ratio and its slope.

Nasrallah et al. (2019) fitted multiple Gaussian functions to a time-series of backscatter intensities (VV, VH and VH/VV) to estimate the date of significant phenology stages for wheat. Wali et al. (2020) explored the sensitivity of temporal backscatter intensities of rice biophysical parameters using a combination of linear regression lines. With this approach, they were able to identify the reproductive growth stages of rice. Schlund and Erasmi (2020) demonstrated the sensitivity of interferometric phase information to estimate the dates of different phenology stages of wheat.

Information about phenological status can increase crop classification accuracy (Bargiel, 2017; Li et al., 2019). However, available studies on crop monitoring using GRD SAR data are mostly limited to the direct use of backscatter intensities and their ratios, along with a few empirical models. These approaches partly utilize the available polarimetric information from dual-pol GRD SAR data. Dual-polarimetric descriptors that characterize different target scattering mechanisms have a wide range of applicability compared to empirical and data-driven models limited to specific crops and regions. In this regard, an unsupervised clustering framework that suitably utilizes the available polarimetric information from dual-pol GRD SAR data is needed to monitor crop growth dynamics.

In general, polarimetric parameters have been directly attributed to the physical properties of the crop canopy (Lopez-Sanchez et al., 2012, 2014; McNairn et al., 2018; Dey et al., 2020b), and has therefore helped monitor crop phenology. Unfortunately, the polarimetric parameters reported in these studies are not immediately apparent in the case of dual-pol GRD SAR data.

The majority of SAR-based crop monitoring studies were limited to small study areas due to the high volume of data processing. For example, the Sentinel-1 constellation acquires data at a rate of approximately 600 GB per day (Ali et al., 2017). This volume of data requires high storage and computational resources for processing. Unfortunately, these resources are limited and restricted for full exploitation to those with access to High-Performance Computing Systems (HPCS). With the advent of cloud platforms such as the Google Earth Engine (GEE) (Gorelick et al., 2017), the NASA Earth Exchange (Nemani et al., 2011), Amazon Web Services (AWS) (Jackson et al., 2010), and Microsoft Azure (Redkar et al., 2009), large-scale remote sensing and geospatial data analysis have become possible with minimum local computational resources (Hird et al., 2017). In this aspect, the web-based GEE platform is designed to make planetary-scale remote sensing data processing manageable and efficient (Gorelick et al., 2017). The free-to-use policy and various in-built GEE algorithms make it an ideal tool for both experts and non-experts alike. The major contributions of the current study are as follows:

- Introduces three new dual-polarimetric descriptors: m_c, θ_c, and H_c.
- Proposes a new unsupervised clustering framework using two parameters (θ_c and H_c) obtained from the dual-pol GRD SAR data.
  - Six feasible clustering zones depicting different scattering mechanisms.
  - Specific to crop monitoring, the proposed clustering framework effectively characterizes different phenological stages.
- Demonstrates how the proposed algorithm is implemented on GEE, making it available for global monitoring with minimal local computational requirements.

The performance of the parameters and clustering framework is analyzed using time-series Sentinel-1 SAR data for monitoring wheat and canola.

2. Study area and dataset

The test site is located near Carman, Manitoba (Canada), covering an intensively cropped area of 26 km $\times$ 48 km. The dominant crops grown in this region include wheat, canola, soybeans, corn and oats, along with a small fraction of acreages in grassland and pasture. The sowing period of crops in this region varies from early to late May, depending on crop variety and cultivation practices. The harvesting period extends until late September.

The nominal size of each field is approximately 800 m × 800 m. Each field comprises 16 sampling locations arranged in two parallel transects separated by 200 m, as shown in Figure 1. During the SMAPVEX-16 campaign, in-situ measurement of vegetation and soil was collected for 50 fields near coincident with satellite acquisitions.

In this study, we have considered 24 fields (13 wheat and 11 canola) for analysis. Figure 1 presents the distribution of the selected fields in the study area. One can find additional details regarding in-situ sampling methods during the SMAPVEX-16 campaign in McNairn et al. (2016); Bhuiyan et al. (2018).

Figures 2 and 3 provide field photos of different growth stages of wheat and canola, respectively. The Manitoba weekly agriculture reports Agriculture (2016) provide additional details regarding crop conditions. Sentinel-1 operates at C-band with a central transmit frequency of 5.405 GHz. In this work, we have utilized the data acquired with the Interferometric Wide swath (IW) mode with a swath width of 250 km. The spatial resolution is 5 m × 20 m in range and azimuth, respectively, and the Noise Equivalent Sigma Zero (NESZ) is -25 dB with the incidence angle varying between 20° to 46°. From the available Sentinel-1 images acquired during the campaign, we have used eight dual-pol (VV-VH) C-band Sentinel-1A GRD SAR in the present study. We have utilized the VV-VH data acquired with IW mode with incidence angle ranging from 30.65° to 41.76°. Complete details of the SAR data utilized in the study are presented in Table 1.

The data were chosen based on the availability of in-situ measurements of crop phenology stages and coincident Sentinel-1A acquisitions for six days of the year (DOY) for wheat (DOY-146, 165, 182, 189, 201, 230) and canola (DOY-146, 165, 182, 189, 206, 225).

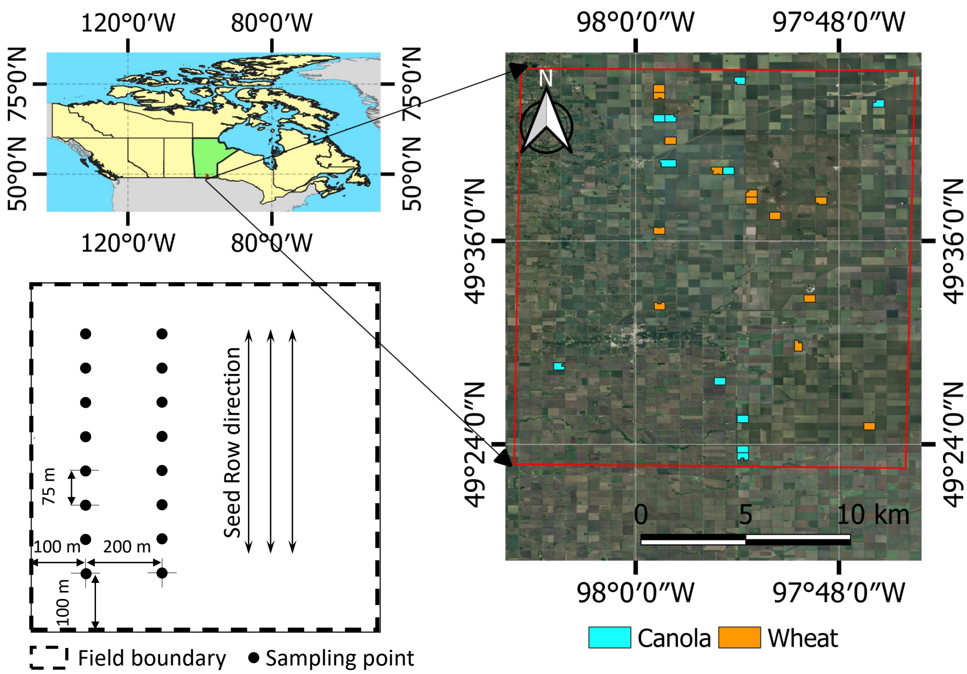


Figure 1: The study area and the distribution of wheat and canola fields in the study area overlaid on a Google earth image. The sampling schema followed for ground truth collection is detailed in the dashed rectangle (bottom left).

3. Methodology

This section proposes three descriptors from the Level-1 S1 GRD SAR data. We express the co-pol purity parameter in terms of the co-pol to cross-pol ratio, which is then used to obtain the scattering-type parameter. The measure of scattering randomness is expressed in terms of the ratio parameter. We utilize these descriptors to introduce a clustering framework

Table 1: Details and specification of Sentinel-1A data used in the present study. Data are acquired from the Carman test site during the SMAPVEX16-MB campaign. The range of incidence angles shown is specific to the location of the sample sites (IW: Interferometric Wide swath)

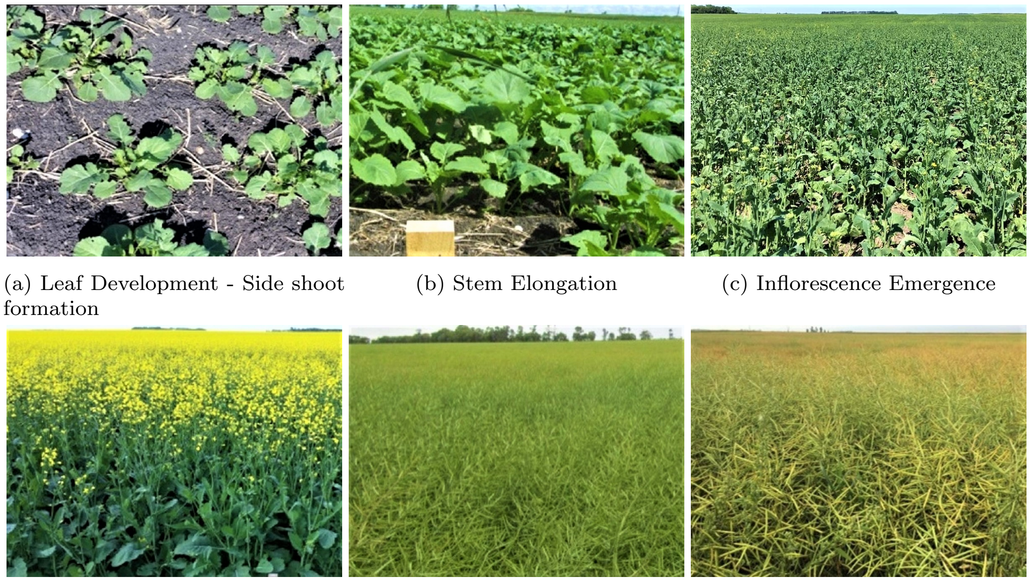
Date	DOY	Acquisition Mode	Incidence angle range (deg.)	Orbit
25-May-16	146	IW	40.18 - 41.76	Ascending
13-Jun-16	165	IW	30.65 - 32.70	Ascending
30-Jun-16	182	IW	40.17 - 41.75	Ascending
07-Jul-16	189	IW	30.64 - 32.69	Ascending
19-Jul-16	201	IW	30.70 - 32.70	Ascending
24-Jul-16	206	IW	40.16 - 41.74	Ascending
12-Aug-16	225	IW	30.65 - 32.70	Ascending
17-Aug-16	230	IW	40.16 - 41.74	Ascending

(a) Leaf Development - Tillering (b) Stem Elongation-Booting (c) Heading

(d) Flowering-Fruit Development (e) Dough formation (f) Maturity

Figure 2: Field photos showing different phenology stages of wheat.

for crop phenology identification. Finally, we present the overall processing chain of the framework using the GEE platform.



(a) Leaf Development - Side shoot formation | (b) Stem Elongation | (c) Inflorescence Emergence | (d) Flowering | (e) Pod Development | (f) Ripening

Figure 3: Field photos showing different phenology stages of canola.

3.1. Dual-polarimetric descriptors

In this section, we present three dual-polarimetric descriptors from the Level-1 S1 GRD SAR data. We propose an unsupervised clustering framework to monitor different crop growth stages based on their diverse scattering characteristics using these descriptors. In the Level-1 S1 GRD SAR data product, we obtain backscatter response either in $(\sigma^{\circ}_{\mathrm{VV}}, \sigma^{\circ}_{\mathrm{VH}})_{\mathrm{dB}}$ or $(\sigma^{\circ}_{\mathrm{HH}}, \sigma^{\circ}_{\mathrm{HV}})_{\mathrm{dB}}$ modes, where H and V are respectively the horizontal and vertical transmit and receive polarization components. The subscript dB represents the GRD SAR data products in decibel (dB) scale. In general, for a monostatic antenna configuration and a natural scene, we assume $\sigma^{\circ}_{\mathrm{XY}} \leq \sigma^{\circ}_{\mathrm{XX}}$ (where X and Y are H or V polarizations respectively) (Cloude, 2009). Using this assumption, we consider the ratio parameter, $0 \leq q = \dfrac{\sigma^{\circ}_{\mathrm{XY}}}{\sigma^{\circ}_{\mathrm{XX}}} \leq 1$, in the linear scale. This parameter has been widely used in the literature as a descriptor for several crop monitoring applications (Della Vecchia et al., 2008; Vreugdenhil et al., 2018; Homayouni et al., 2019). In the GRD product, we do not keep the relative phase information between the XX and XY polarization. Hence, we cannot obtain covariance information from the GRD product, unlike the SLC data. We express the co-pol purity parameter (m_c) in terms of q given in Equation 1. It can be noted that for $q = 1$, $m_c = 0$, and for $q = 0$, $m_c = 1$. In between these two extreme cases, $1 > q > 0$, $0 < m_c < 1$.

$$m_c = \frac{1-q}{1+q}; \quad 0 \leq m_c \leq 1 \tag{1}$$

Utilizing $\sigma^{\circ}_{\mathrm{XX}}$, $\sigma^{\circ}_{\mathrm{XY}}$, and m_c we define two auxiliary quantities as,

$$\tan\zeta_1 = \frac{\sigma^{\circ}_{\mathrm{XX}}}{m_c\, I} \quad \text{and} \quad \tan\zeta_2 = \frac{\sigma^{\circ}_{\mathrm{XY}}}{m_c\, I}, \tag{2}$$

where the total intensity, $I = \sigma^{\circ}_{\mathrm{XX}} + \sigma^{\circ}_{\mathrm{XY}}$. By using a simple relationship, we obtain,

$$\begin{aligned} \tan\theta_c &= \tan(\zeta_1 - \zeta_2) \\ &= \frac{(1-q)^2}{1+q^2-q}; \quad 0^{\circ} \leq \theta_c \leq 45^{\circ} \end{aligned} \tag{3}$$

We can observe from equation (3) that when $m_c = 0$, then $\theta_c = 0^{\circ}$ characterizes complex scattering from targets. Whereas, when $m_c = 1$, then $\theta_c = 45^{\circ}$, characterizes pure scattering from deterministic targets (i.e., trihedral or dihedral). Therefore, the pseudo scattering-type parameter $\theta_c \in [0^{\circ}, 45^{\circ}]$ characterizes different scattering information in between these two cases.

Next, we define the pseudo scattering entropy parameter as,

$$H_c = -\sum_{i=1}^{2} p_i \log_2 p_i; \quad 0 \leq H_c \leq 1 \tag{4}$$

where $p_1 = \dfrac{1}{1+q}$ and $p_2 = \dfrac{q}{1+q}$ are the two pseudo probability measures with $p_1 \geq p_2$. We can observe that $H_c = 1$ for $p_1 = p_2$ (i.e., $q = 1$), whereas $H_c = 0$ for $p_1 = 1$ (i.e., $q = 0$).

Using θ_c and H_c together, we propose an unsupervised clustering framework shown in Fig. 4. The curve (Curve I) represents the unique feasible clustering section in the H_c/θ_c plot. It can be noted that this curve is determined from the theoretical relationship between θ_c and H_c while varying m_c between 0 to 1.

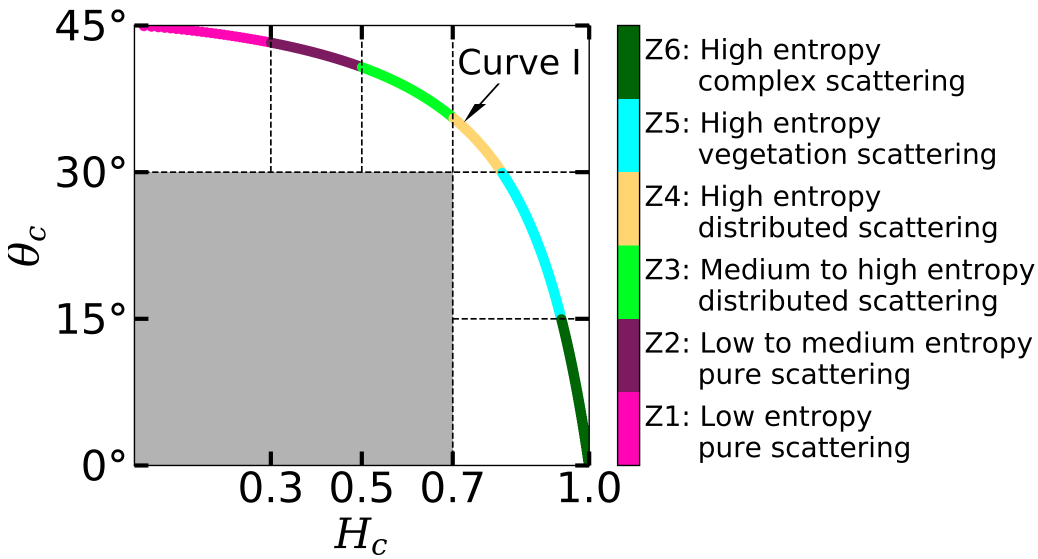


Figure 4: The H_c/θ_c 2D clustering sections. The curve is divided into six zones: Z1 to Z6.

Based on particular scattering characteristics from targets (Cloude and Pottier, 1997), we propose six possible clustering zones: Z1, Z2, Z3, Z4, Z5 and Z6 by splitting H_c into four sub-categories: $[0, 0.3)$, $[0.3, 0.5)$, $[0.5, 0.7)$, and $[0.7, 1.0]$, and θ_c into three sub-categories: $[0°, 15°)$, $[15°, 30°)$, $[30°, 45°]$. Each of these zones represents different scattering phenomena from the scene. The relationship between q and the three proposed descriptors (θ_c, m_c, and H_c) along with the boundary values of the clustering zones is shown in Figure 5 and Table 2.

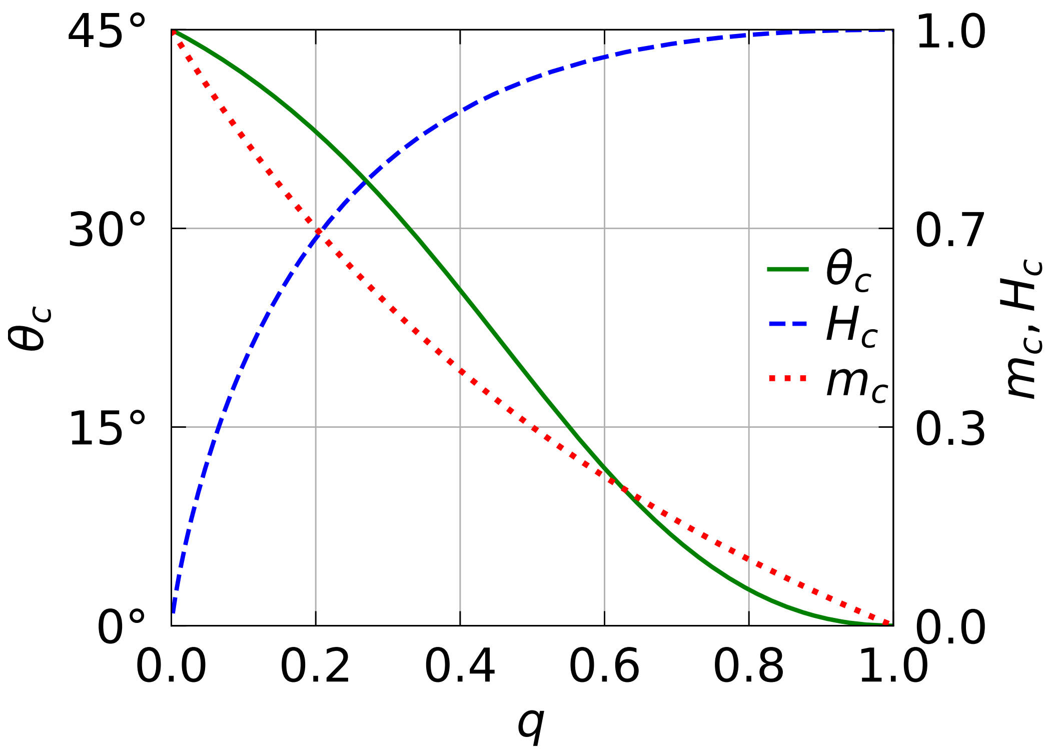


Figure 5: Relationship between the proposed descriptors (θ_c, m_c, H_c) and q.

Cloude (2007) proposed the $H/\overline{\alpha}$ clustering technique for dual-polarimetric (HH-HV or VV-VH) SAR data. Unlike full-polarimetric measurements, the 2×2 covariance matrix is formed using only the column of the scattering matrix to characterize various targets. The scattering angle α is obtained from the eigenvector parameterization, and H is obtained from the eigenvalues of

Table 2: Boundary values of the descriptors (H_c, θ_c, q, and m_c) for adjacent zones in the proposed clustering framework

	H_c	θ_c	q	m_c
—	0	45°	0	1
Z1–Z2	0.30	43.25°	0.06	0.89
Z2–Z3	0.50	40.74°	0.12	0.78
Z3–Z4	0.70	35.59°	0.23	0.62
Z4–Z5	0.81	30.00°	0.33	0.51
Z5–Z6	0.94	15.00°	0.55	0.29
—	1	0°	1	0

the covariance matrix as pseudo probabilities. The average scattering angle $\overline{\alpha}$ is obtained from the two orthogonal polarization states weighted with the two corresponding pseudo probabilities.

It is important to note that we cannot directly apply the $H/\overline{\alpha}$ decomposition technique to characterize target scattering mechanisms for GRD SAR data. Hence, to characterize targets utilizing GRD data, we propose an equivalent scattering angle θ_c based on the approach presented in (Dey et al., 2020a). We present the comparison of the two scattering angles for elementary targets and volume scattering models from a random cloud of anisotropic particles in Table 3. For comparison purpose, the scattering angle θ_c is scaled to the same range of $\overline{\alpha}$ as, $\overline{\theta}_c = 45^\circ - \theta_c$. We can note that all elementary targets reside at the origin, whereas the volume scattering models reside precisely on the lower curve of the $H/\overline{\alpha}$ plane.

Unlike the unsupervised clustering plane formed from the dual-pol H/α framework, the proposed H_c/θ_c framework forms a clustering segment. Both θ_c and H_c are derived from the cross-pol ratio q. However, their physical interpretations for targets are quite different due to their fundamental formula-

Table 3: Comparison between Cloude $\overline{\alpha}$ and $\overline{\theta}_c$ for elementary targets and volume scattering models for dual-polarimetric SAR data.

	Trihedral	Dihedral	Prolate	Oblate	Noise (Identity)
$\overline{\alpha}$	0°	0°	22.5°	10°	45°
$\overline{\theta}_c$	0°	0°	15.3°	4.3°	45°
$H = H_c$	0	0	0.811	0.503	1

tion, even though there is some correlation between the two parameters. On the one hand, the derivation of H_c is equivalent to the von Neumann type of entropy (represented as Shannon entropy) utilizing the pseudo-probabilities in terms of q. On the other hand, the derivation of θ_c follows from the equivalent formalism given in (Dey et al., 2020a). It characterizes scattering-type information using co-pol purity (m_c) and total intensity (I) in terms of q. One can note that their combined use is also supported by a better separation of clusters when the thresholds are defined by any one of these parameters.

The division of the clustering segment is realized from the symmetry relation for the scattering of a polarized wave. The input Stokes vector $\mathbf{S}_i$ and output Stokes vector to the scattering medium $\mathbf{S}_o$ are related by a linear relation of the form: $\mathbf{S}_o = \mathbf{K}\mathbf{S}_i$. Several restrictions are attributed to the Kennaugh matrix $\mathbf{K}$ depending upon the symmetry and reciprocity requirements. Scattering from symmetrical medium makes $\mathbf{K}$ diagonal. In the limit of weak scattering, the linear response of the scattering medium is determined by the ensemble-averaged covariance satisfying the Bethe-Salpeter equation (Cloude and Pottier, 1997). Following some rigorous computation, $\mathbf{S}_o$ can be expressed as a function of the number of scattering events, n. Having specified $\mathbf{S}_o$, we can formulate the expression of the degree of polarization, m

in terms of n. From the definition of entropy S (H_c in this context) given in (Brosseau, 1991; Bicout and Brosseau, 1992), which is a function of the degree of polarization m, satisfying the inequality: $S(m = 1) \leq S \leq S(m = 0)$. Therefore, we observe that S increases with increase in n, as, $S(n = 0) = 0$; $S(n = 1) \approx 0.3$; $S(n = 2) \approx 0.5$, and, $S(n \geq 3) \approx 0.7$, and further increasing n (i.e., higher-order scattering), S saturates for both dual- and full-pol case. Furthermore, for dual cross-pol case, $H_c \approx 0.7$ for randomly oriented cylinders. A similar dependency of the scattering-type parameter (Cloude α) can also be observed as a function of the order of scattering n. We can approximately translate this observation to θ_c.

3.2. Effect of system parameters on the proposed descriptors

In this section, we show the analysis of the effect of polarization combination and frequency. Also, we present a comparative study of conventional dual-pol descriptors from SLC data and the proposed dual-pol descriptors from dual-pol GRD SAR data. In this context, we have utilized the RADARSAT-2 (C-band) and UAVSAR (L-band) data acquired over a Canadian test site for wheat. During the acquisition, wheat was at flowering to heading stage. The sampling points consists of acquisitions from two dates (29 June 2012 and 14 July 2012), and the incidence angle ranges from 22.2° to 26.5°.

We know that longer wavelength SAR signal (L-band) penetration depth is higher than the shorter SAR signal (C-band). Moreover, a shorter wavelength SAR signal (C-band) will suffer relatively more attenuation within vegetation canopies than a longer wavelength SAR signal (L-band). Therefore, we observe the differences in the proposed descriptors for different incident frequencies. From Table 4, we can observe higher values of m_c and lower values of the pseudo scattering entropy in the case of L-band compared to the C-band. It may be due to less attenuation of the L-band compared to C-band. Similarly, we observe that the values of θ_c are more towards a pure scattering-type in the L-band than the C-band in both the dual-pol combinations. Also, we observe some effects of polarization combination on the descriptors. The predominantly vertical structure of the wheat canopy leads to higher interaction of the V-pol than the H-pol. Hence, we observe higher scattering entropy H_c and lower co-pol purity m_c in the VV+VH combination.

Besides, we observed higher values of the scattering-type parameter θ_c in the HH+HV combination than the VV+VH combination. The high value of θ_c indicates that the scattering mechanism is comparatively purer in the HH+HV combination than the VV+VH combination due to less interaction of H-pol with the vertically oriented crop canopy than the V-pol.

Table 4: Effect of frequency on the proposed dual-pol descriptors for wheat.

Frequency	HH+HV			VV+VH		
	m_c	H_c	θ_c	m_c	H_c	θ_c
C-band (5.405 GHz)	0.79 ± 0.08	0.47 ± 0.12	$40.78^\circ \pm 2.30^\circ$	0.53 ± 0.12	0.78 ± 0.1	$30.33^\circ \pm 6.59^\circ$
L-band (1.258 GHz)	0.93 ± 0.03	0.22 ± 0.07	$43.84^\circ \pm 0.59^\circ$	0.83 ± 0.08	0.4 ± 0.14	$41.79^\circ \pm 2.15^\circ$

A comparative study between the conventional SLC dual-pol descriptors: Barakat degree of polarization m, scattering entropy H and Cloude $\overline{\alpha}$, and the proposed GRD dual-pol descriptors: co-pol purity m_c, pseudo scattering-entropy H_c, and pseudo scattering-type parameter θ_c. Table 5 shows the values of the conventional dual-pol descriptors and the proposed descriptors for two dual-pol combinations for the above experiment setup. We observe a negligible difference between H and H_c, and m and m_c. However, we observe a noticeable difference between the scattering-type parameters, $\overline{\alpha}$ and $\overline{\theta}_c$ (kindly note that $\overline{\theta}_c = 45^\circ - \theta_c$). This difference could be because of the parameterization of the eigenvector of the $\mathbf{C}_2$ matrix while deriving the Cloude α. Hence, we can say that our proposed parameters obtained from dual-pol GRD SAR data possess equivalent information as the conventional parameters derived from dual-pol SLC data.

Table 5: Comparison of conventional dual-pol descriptors from SLC data and the proposed dual-pol descriptors from dual-pol GRD SAR data.

Channels	SLC			GRD		
	m	H	$\overline{\alpha}$	m_c	H_c	$\overline{\theta_c}$
HH+HV	0.80 ± 0.08	0.45 ± 0.15	$12.23^\circ \pm 4.24^\circ$	0.79 ± 0.08	0.47 ± 0.12	$4.22^\circ \pm 2.3^\circ$
VV+VH	0.55 ± 0.11	0.76 ± 0.10	$24.27^\circ \pm 5.45^\circ$	0.53 ± 0.13	0.78 ± 0.10	$14.67^\circ \pm 6.59^\circ$

3.3. Sentinel-1 dual-pol descriptors in GEE

This section describes the extraction process of the proposed polarimetric descriptors from the Sentinel-1 dual-pol GRD SAR data on the GEE platform. The overall processing framework comprises three major blocks: data preparation, clustering and temporal analysis, as shown in Figure 6.

In the data preparation block, we import the Level-1 Ground Range Detected (GRD) Sentinel-1 backscatter coefficient (i.e., σ° in decibel) data into the GEE platform. The imported temporal data stack is cloud filtered using three filters:

- Metadata filter (bands: VV, VH, incidence angle, instrument mode: IW, and orbit: ascending)
- Temporal filter (date range: 25 May 2016 to 17 August 2016)

- Spatial bound filter (region of interest: shapefile of the study area)

Subsequently, we use two masks to generate a valid pixel data stack. As described in section 3.1, the first mask (i.e., $\sigma^{\circ}_{\mathrm{VV}} > \sigma^{\circ}_{\mathrm{VH}}$) ensures estimation constraints of the proposed descriptors, whereas the second mask (i.e., $\sigma^{\circ}_{\mathrm{VV}} > -20\,\mathrm{dB}$) separates out water bodies. The backscatter values in the valid pixel data stack are then converted into a linear scale. Further, we use a 5×5 boxcar filter to despeckle the data.

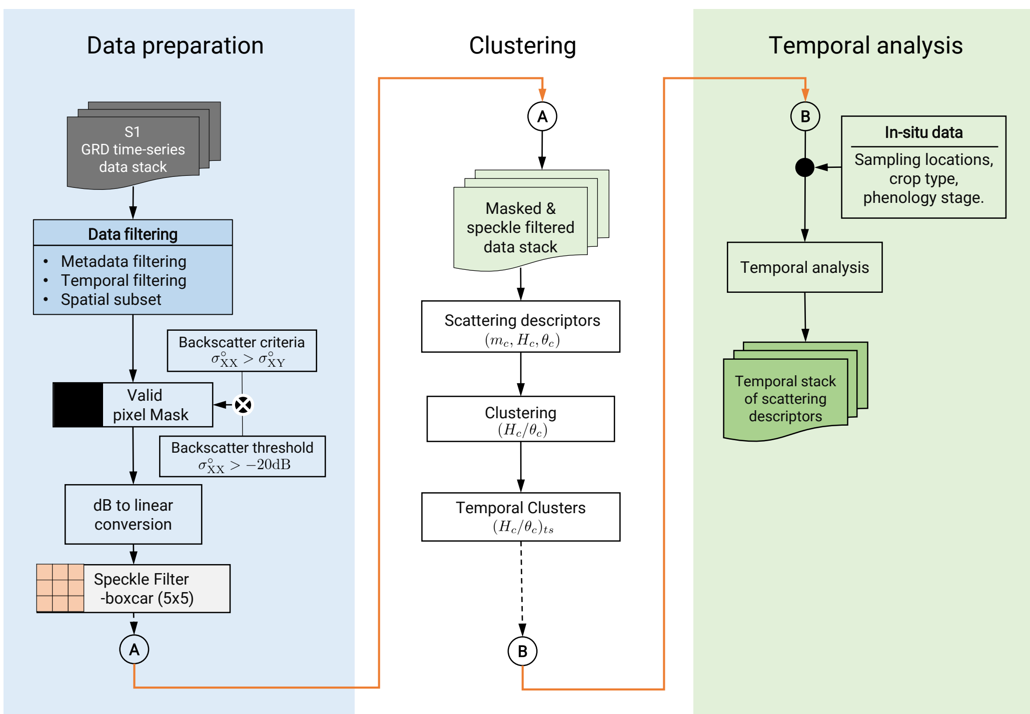


Figure 6: The proposed schematic workflow to derive the dual-polarimetric descriptors from Sentinel-1 dual-pol GRD SAR data on the GEE platform.

In the clustering block, we first generate the dual-pol descriptors using the Equations 1- 4 from the valid and speckle filtered data stack. Further, we utilize these descriptors to generate the H_c/θ_c clusters for each scene in the temporal stack. The temporal analysis block starts by importing in-situ data such as sampling locations, crop type and crop phenology stages. Subsequently, we utilize these data to analyze the temporal stack of dual-pol descriptors and the H_c/θ_c clusters. Further, to complement the analysis, we generate temporal maps of each descriptor over the study area.

4. Results and discussion

In this section, we analyze the temporal dynamics of crops using the proposed dual-polarimetric descriptors. Furthermore, we utilize the proposed clustering framework to assess the phenological stages of the two crops (wheat, canola) from the C-band Sentinel-1 dual-pol GRD SAR data. The description of the phenological stages for wheat and canola are presented in Tables 6 and 7 respectively.

Table 6: Phenology stages of wheat. The BBCH (**B**iologische **B**undesanstalt, Bundessortenamt und **CH**emische Industrie) codes of each phenology stage are also highlighted.

Phenology stage	BBCH code	Description
Leaf development	10-19	1-9 or more leaves unfolded
Tillering	20-29	Formation of 1-9 or more number of tillers
Stem elongation	30-39	Elongation of first internode to fully unrolled flag leaf
Booting	41-49	Flag leaf sheath extended to first awns visible
Heading	51-59	First spikelet to completely emerged heads
Flowering-fruit development	61-77	Beginning of flowering and formation of grains with milk
Doughstage	83-89	Development of soft to Hard dough
Maturity	92-97	Grain turns very hard and over ripened; grain loosening in day-time

4.1. Temporal dynamics of the dual-polarimetric descriptors

In this section, we present the temporal analysis of m_c, θ_c and H_c across the phenological stages of wheat and canola over Carman, Manitoba, Canada. The spatio-temporal changes of m_c, θ_c and H_c are shown in Figure 7, Figure 8, and Figure 9, respectively, over the entire test site. Variations for

Table 7: Phenology stages of canola. The BBCH (**B**iologische **B**undesanstalt, Bundessortenamt und **CH**emische Industrie) codes of each phenology stage are also highlighted.

Phenology stage	BBCH code	Description
Leaf development	10-19	1-9 or more leaves unfolded
Side shoot formation	20-29	Formation of 1-9 or more side shoots
Stem elongation	30-39	Formation of 1-9 or more extended internodes
Inflorescence emergence	50-59	Formation of flower buds, still enclosed by leaves
Flowering	60-69	Starting from first flower opening to the majority of petals fallen
Pod development	71-79	Formation of pods and reaching their full size
Ripening	80-89	Green seeds hardens and turns into dark

all three parameters are evident with crop growth starting from early leaf development to maturity for most agricultural fields. We also present the temporal dynamics of H_c/θ_c clusters as shown in Figure 10, to assess the crop growth condition.

On DOY-146, most wheat and canola fields show high values of m_c (wheat: 0.81 ± 0.08 and canola: 0.82 ± 0.08), θ_c (wheat: $41.37^\circ \pm 2.04^\circ$ and canola: $41.6^\circ \pm 1.91^\circ$) and medium to low value of H_c (wheat: 0.43 ± 0.12 and canola: 0.42 ± 0.13). These responses are due to the minimal crop cover, before significant vegetative growth and leaf development. Hence, the soil characteristics (i.e., moisture and surface roughness) dominate the backscatter response. Therefore, the effect of soil roughness on the backscatter response is significant (Wiseman et al., 2014), which may have led these sample pixels to cluster in the low to medium entropy pure scattering zones (viz., Z1, Z2) in the H_c/θ_c map (Figure 10).

With crop growth advancing to the inflorescence stage, we observe an overall decrease in the values of m_c, which is evident in Figure 7. Thus, on DOY-206, we observe medium to low values of m_c (wheat: 0.43 ± 0.07 and canola: 0.42 ± 0.1). From the flowering to maturity stage, the canopy

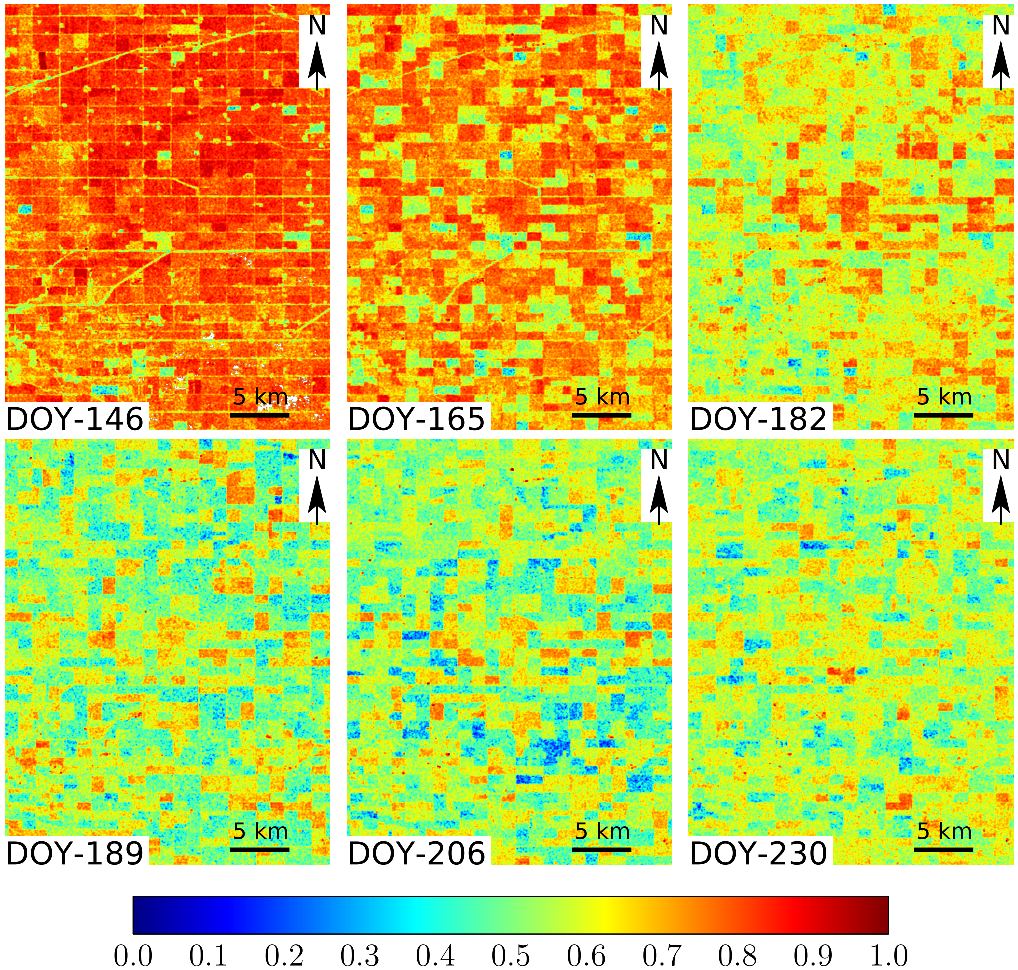


Figure 7: Temporal variation of m_c over the study area.

density increases as crop biomass increases (Wiseman et al., 2014; Hariharan et al., 2018). Therefore, as reported in (Sarabandi, 1991; Wang et al., 2019; Ratha et al., 2019), we also observe similar high scattering randomness at this stage.

Moreover, during this period, the observed backscatter response is expected to be dominated by the upper canopy layer. Additionally, the values

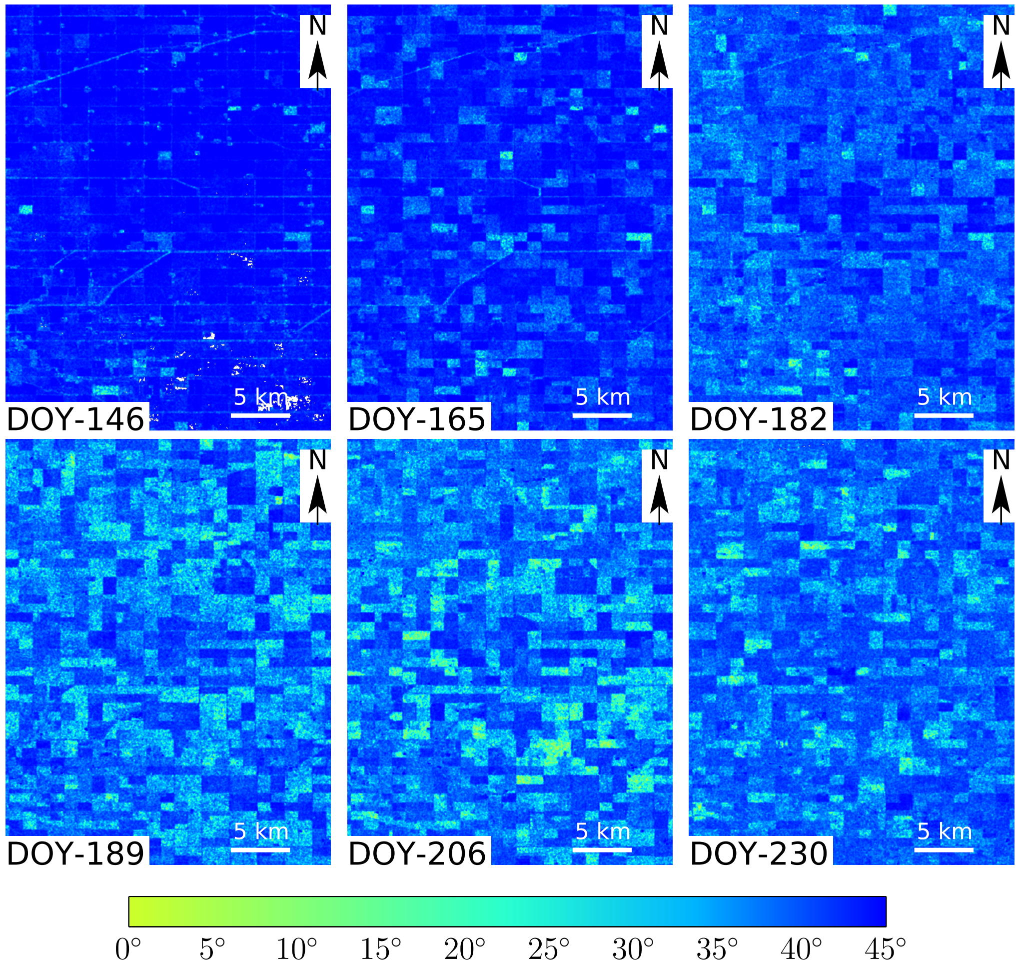


Figure 8: Temporal variation of pseudo scattering type parameter (θ_c) over the study area.

of θ_c for wheat are 24.51° ± 6.34°, and for canola 25.23° ± 4.7° (Figure 8). These values are indicative of low random scattering within the resolution cells. We also observe an increasing trend of H_c due to the randomly oriented canopy structure. The values of H_c for wheat and canola are 0.86 ± 0.06 and 0.86 ± 0.05, respectively (Figure 9). Due to randomness in the vegetation structure on DOY-206, we observe dominance within the high entropy vegetation scattering zone (Z5) in Figure 10.

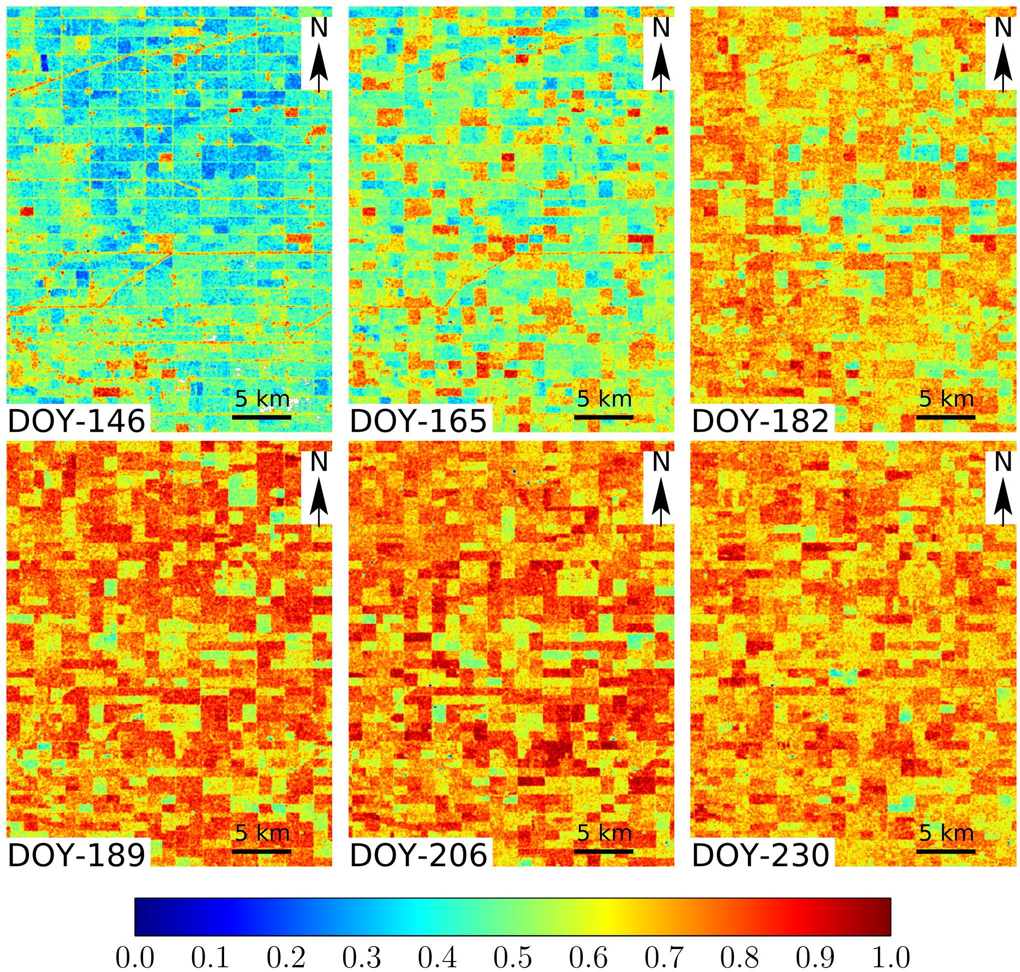


Figure 9: Temporal variation of pseudo scattering entropy parameter (H_c) over the study area.

One can note that all three polarimetric descriptors show a trend reversal at early crop senescence. This change could be due to the randomness variation corresponding to morphology attributes with a likely decrease in canopy moisture. We can observe from Figure 7 that m_c significantly increases during the ripening stage (DOY-230). At this stage, the values of m_c for wheat and canola are 0.52 ± 0.08 and 0.55 ± 0.08, respectively. A similar trend is

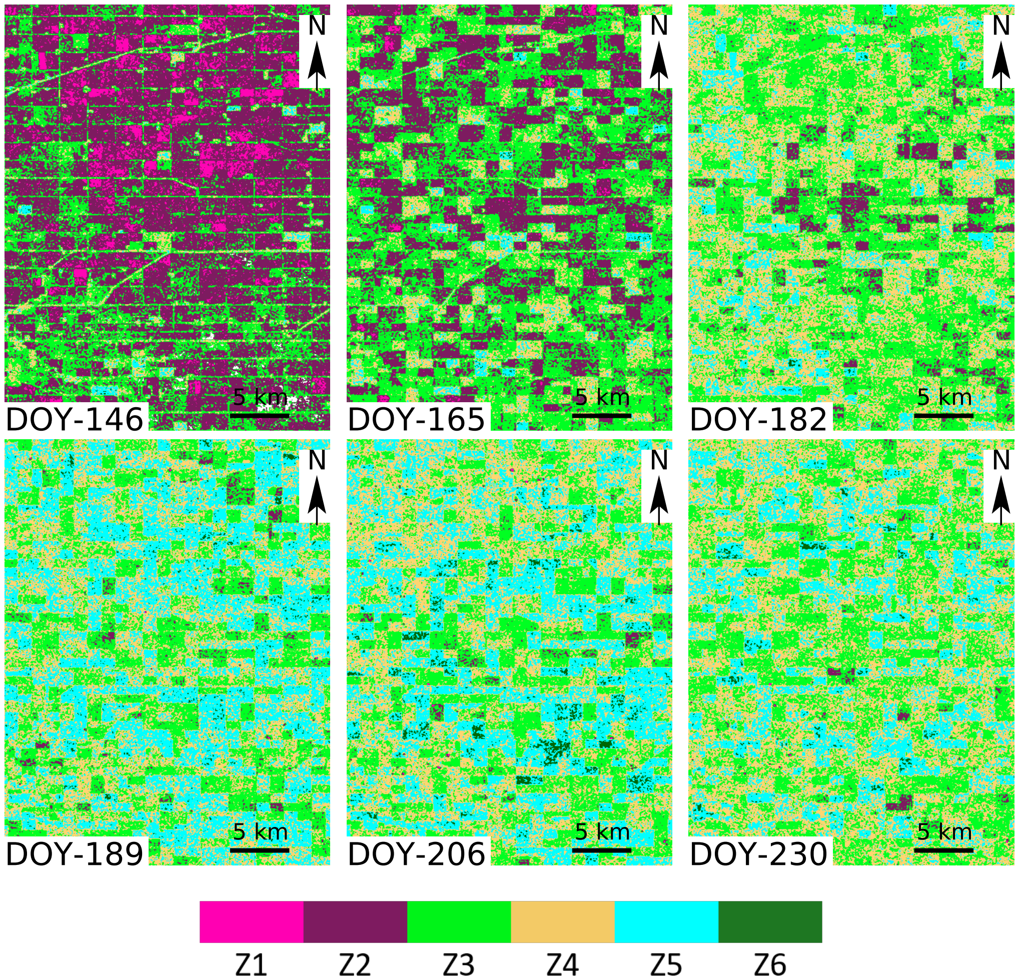


Figure 10: Temporal variation of H_c/θ_c clusters over the study area.

also observed for θ_c and H_c from Figures 8 and Figure 9. These notable changes in the polarimetric descriptors might be due to the enhanced ability of the radar wave to penetrate into the moderately dry crop canopy. Also, the vegetation water content variations might have decreased the SAR signal attenuation within the resolution cell.

4.2. Analysis over sampling fields

In the following sections, we provide a detailed quantitative analysis of the three descriptors (m_c, θ_c, and H_c) and the novel clustering framework for wheat and canola. In this study, we considered a total of 24 sample fields (wheat: 13 and canola: 11) for sensitivity and performance evaluation of the descriptors during temporal morphological changes in the canopies.

4.2.1. Wheat

First, we analyse temporal characteristics of m_c, H_c and θ_c for different phenological stages of wheat. We considered a total of 48 sampling points in three different fields (Field no. 62, 220, and 233) for assessing the temporal dynamics of θ_c, H_c and m_c. We also evaluate temporal variations of the H_c/θ_c clusters according to wheat phenology.

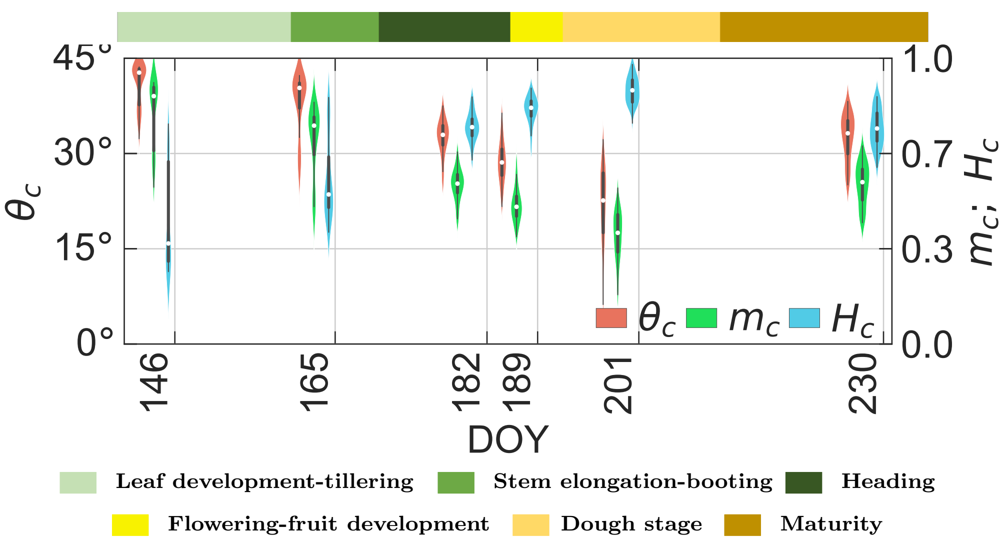


Figure 11: Temporal variation of m_c, H_c and θ_c for the growth stages of wheat. The white dot represents the median value, the black bar in the center represents the standard boxplot. On either side of the boxplot is a kernel density estimation displaying the shape of the data distribution.

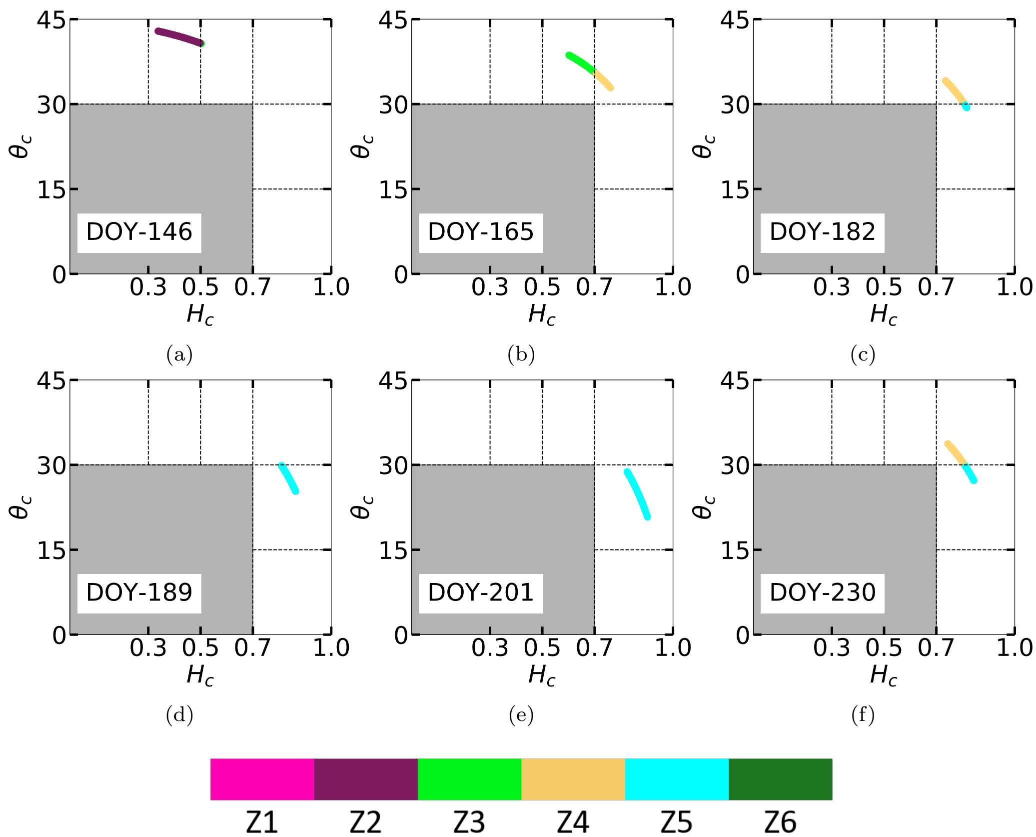


Figure 12: Temporal dynamics of the H_c/θ_c data cluster for wheat during entire growth period.

We have shown the temporal dynamics of the parameters m_c, H_c and θ_c using a series of standard violin plots. The width of the violin represents the probability that the sampling points portion will take on the given value. We can observe from Figure 11 that m_c, H_c and θ_c are sensitive to wheat morphological changes. For example, we note that during the early leaf development stage (around DOY-146), m_c and θ_c show differential scattering information due to the presence of a minimal crop canopy. However, on DOY-201, wheat has advanced to the dough stage and consequently, the interaction of the radar wave with the matured canopy structure has increased. Therefore at this stage, the increase in the cross-pol component has decreased the value of m_c.

On DOY-146, high values of m_c and θ_c are evident from Figure 11, reaching 0.8 ± 0.12 and $40.82° \pm 3.41°$, respectively. The distributions of m_c and θ_c are left-skewed, indicating that most samples fall towards higher values of m_c and θ_c. These high values correspond to pure scattering, which is due to the dominance of the soil contribution relative to vegetation. As a result, most data points are clustered in the low entropy pure scattering zone (Z2) (Figure 12a). However, it is noteworthy that a fraction of the backscatter response originates from the crop leaves and side tillers. The interaction of radar waves with these structures has produced a small cluster with $\approx$11 % of data points in the medium entropy zone (Z3) (Table. 8).

We note that a few fields have advanced to the tillering stage during this period due to early sowing. Hence, new tillers in those fields might have decreased m_c and increased H_c. During stem elongation and booting, on DOY-165, we observe a decrease in m_c and θ_c. The values of m_c and θ_c are 0.71 ± 0.12 and $37.87° \pm 4.88°$, respectively. This decrease is likely due to changes in crop morphology in the vertical direction, with an increase in the main stem and side tillers (Figure 2b). The distributions of m_c and θ_c have shifted towards lower values. However, depending on the difference in the growth pattern, bi-modal distributions of m_c and θ_c are observed.

During this period, high crop canopy density enhances scattering entropy, shifting the data points towards medium to high entropy zones (Z3, and Z4) as shown in Figure 12b. Hence, 57.7 % of the data points are clustered within the medium entropy zone, while 42.3 % are clustered within the high entropy zone (Table 8). Subsequently, during the heading stage (DOY-182), we observe a considerable drop in the mean values of m_c and θ_c from Figure 11. These values decrease by 21.13 % and 14.21 %, respectively, when compared to the previous date.

On DOY-182, the crops are in their advanced vegetative stage. The radar response is similar for all fields due to their comparable scattering randomness. The standard deviations of the sample distributions have decreased significantly. During this period, we observe from Figure 12c a shift in the data clusters from the medium entropy (Z3) to the high entropy zone (Z4). This shift could be due to changes in the wheat canopy structure during the heading stage. At this stage, the distribution of plant biomass shifts towards the upper layer of the canopy. Thus, a major contribution of scattering is from the upper canopy layer. A small proportion of data points ($\approx$15 %) are clustered in the vegetation scattering zone (Z5) (Table. 8), which might be due to early flowering of these wheat fields.

With the advancement of wheat phenology to flowering stage (Figure 2d) on DOY-189, we observe a further drop in m_c (0.49 ± 0.06) and θ_c ($28.72^\circ \pm 3.39^\circ$). During this period, the wheat canopy forms a complex structure due to the appearance of flowers on the upper portion of the canopy layer. Interestingly, randomness in scattering during the flowering stage is more noticeable in the distribution of θ_c values. Moreover, the tail of the distribution becomes more comprehensive than the previous growth stage (Figure 11). The spread in the distribution of θ_c indicates multiple scattering mechanisms. Moreover, the shape of the distribution for m_c is almost equivalent to θ_c, with an overall shift towards lower values. An increase in pseudo entropy has displaced the H_c/θ_c cluster towards the high entropy vegetation scattering zone (Z5) as shown in Figure 12d.

Dough and maturity stages continued from late July (DOY-201) to mid-August (DOY-225) with the values of m_c, and θ_c reaching their minimum when the crop advanced from flowering to early dough on DOY-201. The mean values of m_c and θ_c reach 0.38 ± 0.09 and $21.56° \pm 6.24°$ respectively. We observe a broad spread of θ_c values in Figure 11, which may be due to randomly oriented wheat stems and heads (Figure 2e). Wu et al. (1985) reported a similar phenomenon given that during the heading stage, a significant portion of the total scattering occurs from the wheat heads. It is noteworthy that the distributions are bi-modal, denoting two primary scattering sources: the thick upper canopy layer and the relatively less dense bottom canopy. We observe that all data points cluster in the high entropy and vegetation scattering zone (Z5) in the H_c/θ_c plot (Figure 12e).

Table 8: Temporal variation in the percentage of data points in each zone for different phenology stages of wheat. The zone with the maximum number of points at a particular phenology stage is highlighted in bold. Each row represents a phenology stage, and the solid line in between two phenology stages represents a significant variation in the temporal trend for the zones.

DOY	Z1	Z2	Z3	Z4	Z5	Z6	Growth stage
146	0.0	**89.4**	10.6	0.0	0.0	0.0	Leaf Development-Tillering
165	0.0	0.0	**57.7**	42.3	0.0	0.0	Stem Elongation-Booting
182	0.0	0.0	0.0	**84.6**	15.4	0.0	Heading
189	0.0	0.0	0.0	0.0	**100**	0.0	Flowering-Fruit development
201	0.0	0.0	0.0	0.0	**100**	0.0	Dough stage
230	0.0	0.0	0.0	**65.4**	34.6	0.0	Maturity

During the ripening stage, the canopy moisture content drops rapidly. As a result, penetration of the SAR signal into the crop canopy increases and hence, there is a substantial scattering contribution from the ground to the total backscatter. A trend reversal is observed for all the dual - polarimetric descriptors when the crop reaches the early mature stage (DOY-230). During this period, the mean values of m_c and θ_c increase to 0.56 ± 0.07 and $32.54° \pm 3.65°$, respectively. The median of the distribution shifts towards higher values (Figure 12f). We observe a decrease in the spread of distributions for both m_c and θ_c in Figure 11 indicating uniformity in the scattering mechanism.

A decrease in scattering entropy shifts the H_c/θ_c cluster towards the distributed scattering zone (Z4) from random scattering, as shown in Figure 12f. However, 34.6 % of the data points are clustered within the vegetation scattering zone (Z5) with 65.4 % of the data points clustered in Z4 (Table 8). The Z5 clusters appearance might be due to the late maturity of wheat, which is also in agreement with the bi-modal distribution of the m_c and θ_c parameters in Figure 11.

The proportion of data points over different scattering regions for other crop phenological stages is presented in Table 8. The results indicate a smooth transition of scattering mechanisms throughout the growing cycle of wheat. Using the proposed scattering descriptors and the novel clustering framework, we capture different scattering mechanisms at each wheat growth stage.

4.2.2. Canola

This section analyses the temporal characteristics of m_c, H_c and θ_c for different phenological stages of canola. In total, 48 sampling points in three canola fields (Field no. 206, 208, and 224) are used to assess the temporal dynamics of these parameters. We also evaluate the temporal variations of the H_c/θ_c cluster according to canola phenology.

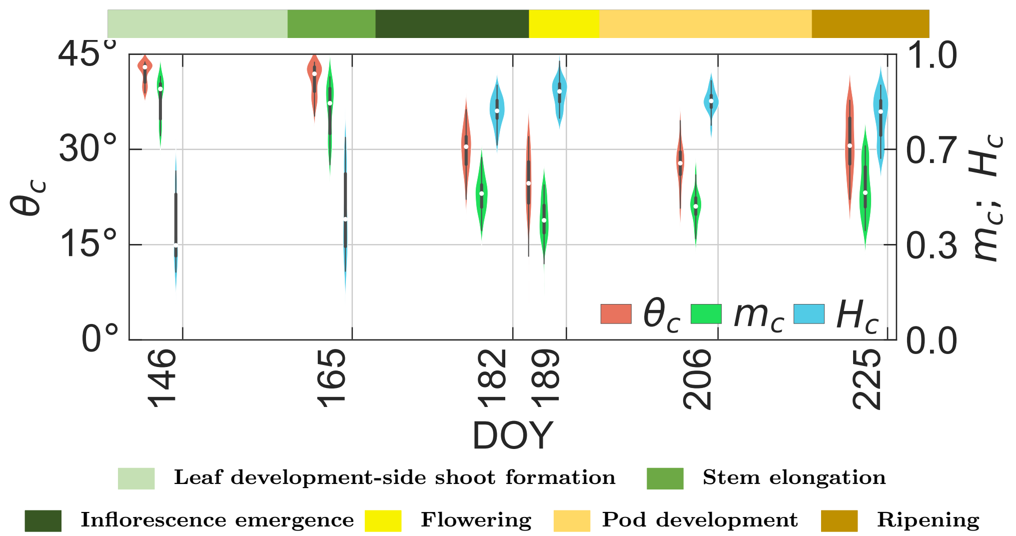


Figure 13: Temporal variation of m_c, H_c and θ_c for the growth stages of canola. The white dot represents the median value, the black bar in the center represents the standard boxplot. On either sides of boxplot is a kernel density estimation to show the distribution shape of the data.

Canola is a broadleaf crop with a distinctive canopy structure at every growth stage (McNairn et al., 2018; Mandal et al., 2020b). The seeding of the canola crop was completed by mid-May, as indicated in the in-situ data. Until the beginning of June, the plant advanced to its vegetative growth stage. The plant develops a dense rosette of leaves near the soil surface during the leaf development, as evident from Figure 3a. However, the size of these leaves is comparable to the wavelength of the C-band ($\approx 5.6\,\mathrm{cm}$).

On DOY-146, the mean value of $\theta_c \approx 40^{\circ}$ and $m_c \approx 0.8$ which indicate dominant scattering from exposed soil due to sparse vegetation cover (Figure 13). In Figure 14a, we observe that a majority of data points are clustered in zone Z2, which is characterized by medium entropy pure scattering. A few data points are in the medium to high entropy distributed scattering zone (Z3) as this crop advances to the leaf development stage.

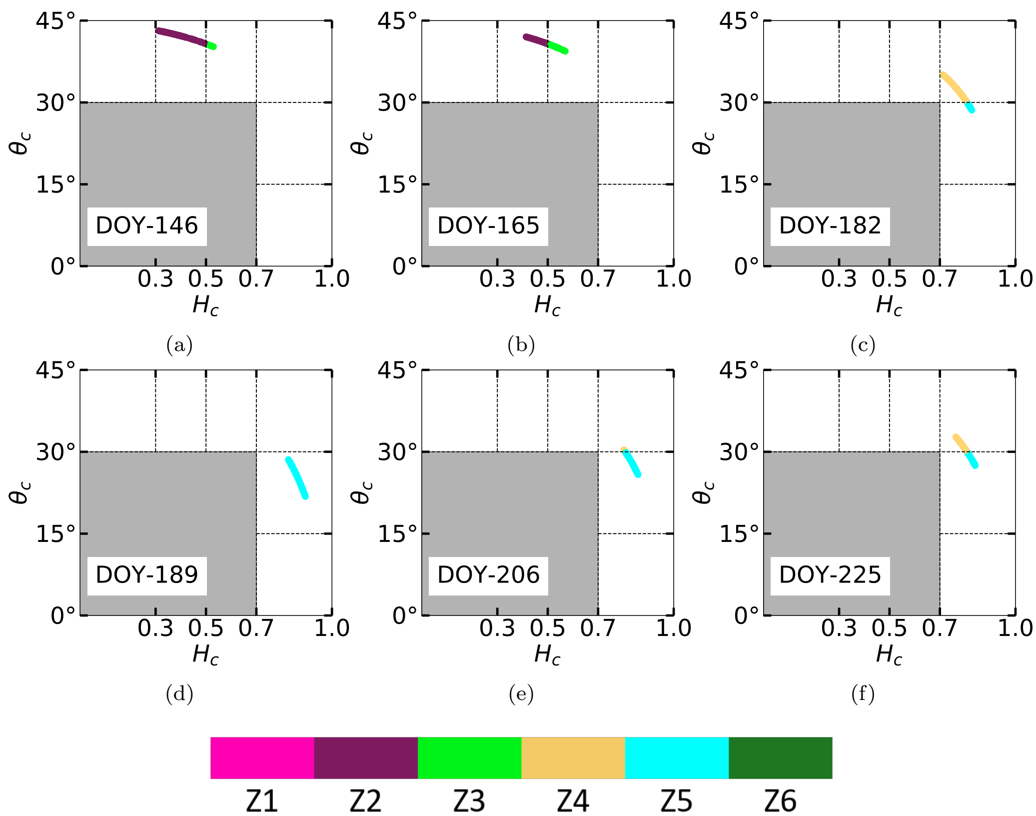


Figure 14: Temporal dynamics of the H_c/θ_c data cluster for canola during entire growth period.

As growth progresses to stem elongation, a noteworthy change into a vertical plant structure can be observed from Figure 3b. At this stage, the attenuation of V-polarized waves increases. Further, an increase in biomass and PAI due to increased leaf density and branch formation leads to increased scattering randomness. On DOY-165, the H_c/θ_c cluster shifts from the pure (Z2) to the distributed scattering zone (Z3) due to a substantial amount of m_c component during stem elongation. The accumulation of points has increased in the distributed scattering zone (Z3) due to the matured crop morphology.

During inflorescence emergence (Figure 3c), flower buds develop and leaf density increases significantly. Consequently, we observe a change in the data cluster on DOY-182 (Figure 14c). Further, the values of m_c and θ_c dropped to 0.51±0.06 and 29.94°±3.38°, respectively. On the other hand, the formation of branches increases scattering entropy. Hence, a shift in the H_c/θ_c cluster from medium entropy zone (Z3) to high entropy zone (Z4) is evident on DOY-182 (Figure. 14c).

During the flowering stage, the buds develop into flowers, and the main stem and branches grow (Figure 3d). On DOY-189, the scattering mechanism of all the data points is shifted towards high entropy vegetation scattering zone (Z5) (Figure 14d). We may attribute this shift to the development of a complex canopy geometry during the flowering and early pod development stage. As pods form, the canopy drops leaves. The decline of leaf cover followed by the development of pods dramatically changes the crop geometry. The canopy architecture becomes more random, with pods creating needle-like structures oriented randomly.

During the development of pods canola develops a dense, random canopy structure. Hence, the mean values of m_c (0.47±0.06) and θ_c (27.6°±3.36°) are minimum on DOY-206. A majority of data points (≈ 91 %) are clustered into the high entropy vegetation scattering zone (Z5). At the same time, a small percentage (≈9 %) of data points are clustered into Z4 (Figure 14e). This small cluster may reflect the change in crop morphology as leaf area declines and the SAR signal interacts more with the needle-like canopy (Figure 3e).

Table 9: Temporal variation in the percentage of data points in each zone for different phenology stages of canola. The zone with the maximum number of points at a particular phenology stage is highlighted in bold. Each row represents a phenology stage and the solid line in between two phenology stages represents a significant variation in the temporal trend for the zones.

DOY	**Z1**	**Z2**	**Z3**	**Z4**	**Z5**	**Z6**	**Growth stage**
146	0.0	**89.8**	10.2	0.0	0.0	0.0	Leaf Development Side shoot formation
165	0.0	48.9	**51.1**	0.0	0.0	0.0	Stem Elongation
182	0.0	0.0	0.0	**84.1**	15.9	0.0	Inflorescence Emergence
189	0.0	0.0	0.0	0.0	**100**	0.0	Flowering
206	0.0	0.0	0.0	9.1	**90.9**	0.0	Pod Development
225	0.0	0.0	0.0	**54.5**	45.5	0.0	Ripening

Subsequently on DOY-225, we observe an increase in m_c and θ_c values to 0.53±0.09 and 30.8°±4.63°, respectively. This increase in the values of the descriptors might be due to the decrease in overall canopy moisture content at maturity. As canopy moisture declines, the SAR signal can penetrate deeper into the crop canopy. Hence, at late maturity, there might be a greater contribution from the soil. Because of these physical changes, the H_c/θ_c

cluster shifts towards the distributed scattering zone (Z4) (Figure 14f). The bi-modal distribution of the parameters m_c and θ_c (Figure 13) indicates two major sources of scattering. In particular, some fraction of the crop may be entering the mid/end-ripening stage, resulting in higher values of m_c and θ_c. In contrast, other canopies may be just entering early ripening, resulting in lower values. We observe this difference in Figure 14f. Notably, 54.5 % of the data points fall in distributed scattering zone (Z4), whereas 45.5 % of the data points are in the vegetation scattering zone (Z5). The proportion of data points over different scattering regions for all phenological stages of canola is presented in Table 9. The results indicate a smooth transition of scattering mechanisms throughout the growing cycle. Consequently, the proposed descriptors exhibit high sensitivity to the phenological stages of both wheat and canola. Hence, these descriptors are useful in monitoring phenological changes for both crops.

5. Conclusion

In this study, we propose three polarimetric descriptors from dual-pol Sentinel-1 (S1) GRD SAR data. These parameter are: the pseudo scattering-type parameter (θ_c), the co-pol purity parameter (m_c), and the pseudo scattering entropy parameter (H_c). We have expressed these descriptors in terms of $q = \sigma^{\circ}_{\mathrm{XY}}/\sigma^{\circ}_{\mathrm{XX}}$, with $0 \leq q \leq 1$. Additionally, we have proposed a novel unsupervised clustering framework using θ_c and H_c.

We have used the dual-pol descriptors and the novel clustering framework to analyze temporal growth dynamics of wheat and canola over a Canadian test site. The results are very encouraging in assessing crop dynamics for different major phenological stages. The high sensitivity of these descriptors to different crop growth stages is evident in this context.

In the scope of this study, we have characterized diverse crop phenological stages in terms of the physical scattering of the electromagnetic wave from targets using the GRD SAR data. The unsupervised clustering framework using H_c/θ_c contains six zones representing different physical scattering mechanisms. Hence, it provides essential information about the crop growth stages without any *a priori* knowledge and therefore very useful in interpreting the available radar data.

The temporal analysis of the proposed descriptors revealed their high sensitivity across different phenology stages of wheat and canola. The dynamic range of θ_c from leaf development to fruit development of wheat is 41° to 21.6°. Similarly, the variations of m_c and H_c are 0.8 to 0.38 and 0.35 to 0.9, respectively. Similar dynamic ranges of these parameters are also evident for canola crop. Furthermore, the proposed clustering schema efficiently captured the diverse phenology stage of wheat and canola. For leaf development and tillering stages of wheat, 90 % of the sampling points are clustered into low to medium entropy pure scattering zone. During flowering and fruit development stages, 100 % of the sampling points are shifted into high entropy vegetation scattering zone. Subsequently, during the maturity stage, the clusters of the sampling points were split between the high entropy vegetation scattering and high entropy distributed scattering zones with $> 65\,\%$ of the sampling points in the high entropy distributed scattering zone. Similar cluster dynamics are observed for the canola crop.

However, it is computationally intensive to implement these algorithms for a high volume of temporal data from a global agricultural monitoring perspective. To overcome this limitation, we utilized the cloud-based platform (GEE) to acquire and process the dense time-series data of Sentinel-1. Implementing the algorithms in GEE also facilitates efficient generation of global maps of crop phenology stages.

This study only used the GRD SAR data product to formulate the target characterizing descriptors demonstrating promising results for natural targets. We can further extend this study to different crop types and different dual-pol SAR sensors configurations. The proposed descriptors should be beneficial in studying natural ecosystems with upcoming dual-pol NASA-ISRO Synthetic Aperture Radar Mission (NISAR) and Sentinel SAR constellation.

Acknowledgement

The authors are grateful to the SMAPVEX16 science team for providing ground truth information. The authors would like to thank the Google Earth Engine team for providing the free SAR data processing platform. Authors also acknowledge the GEO-AWS Earth Observation Cloud Credits Program, which supported the computation with Sentinel-1 on AWS cloud platform through the project: "AWS4AgriSAR-Crop inventory mapping from SAR data on a cloud computing platform", and formed the testbed for processing pipelines. Mr. Narayanarao B. and Mr. Subhadip Dey would like to acknowledge the support from the Ministry of Education (formerly Ministry of Human Resource and Development-MHRD), Govt. of India, towards their doctoral research work. The authors want to thank the support of the Spanish Ministry of Science and Innovation, State Research Agency (AEI) and the European Regional Development Fund under project TEC2017-85244-C2-1-P. The authors are thankful to the overleaf team (https://overleaf.com/) for providing the latex editing platform.

References

Agriculture, M., B., 2016. Agriculture Province of Manitoba. URL: https://www.gov.mb.ca/agriculture/crops/seasonal-reports/crop-report-archive/index.html.

Ainsworth, T., Kelly, J., Lee, J.S., 2009. Classification comparisons between dual-pol, compact polarimetric and quad-pol SAR imagery. ISPRS Journal of Photogrammetry and Remote Sensing 64, 464–471.

Ainsworth, T.L., Kelly, J., Lee, J.S., 2008. Polarimetric analysis of dual polarimetric SAR imagery, in: 7th European Conference on Synthetic Aperture Radar, VDE. pp. 1–4.

Ali, I., Naeimi, V., Cao, S., Elefante, S., Bauer-Marschallinger, B., Wagner, W., 2017. Sentinel-1 data cube exploitation: Tools, products, services and quality control, in: Proc. Big Data Space, pp. 40–43.

Arias, M., Campo-Bescós, M.Á., Álvarez-Mozos, J., 2020. Crop classification based on temporal signatures of sentinel-1 observations over navarre province, Spain. Remote Sensing 12, 278.

Bargiel, D., 2017. A new method for crop classification combining time series of radar images and crop phenology information. Remote sensing of environment 198, 369–383.

Bhuiyan, H.A., McNairn, H., Powers, J., Friesen, M., Pacheco, A., Jackson, T.J., Cosh, M.H., Colliander, A., Berg, A., Rowlandson, T., et al., 2018. Assessing SMAP soil moisture scaling and retrieval in the Carman (Canada) study site. Vadose Zone Journal 17, 1–14.

Bicout, D., Brosseau, C., 1992. Multiply scattered waves through a spatially random medium: entropy production and depolarization. Journal de Physique I 2, 2047–2063.

Bousbih, S., Zribi, M., Lili-Chabaane, Z., Baghdadi, N., El Hajj, M., Gao, Q., Mougenot, B., 2017. Potential of Sentinel-1 radar data for the assessment of soil and cereal cover parameters. Sensors 17, 2617.

Brisco, B., Brown, R., Gairns, J., Snider, B., 1992. Temporal ground-based scatterometer observations of crops in western Canada. Canadian journal of remote sensing 18, 14–21.

Brosseau, C., 1991. Polarization transfer and entropy transformation. Optik (Stuttgart) 88, 109–117.

Cloude, S., 2007. The dual polarization entropy/alpha decomposition: A palsar case study. ESASP 644, 2.

Cloude, S., 2009. Polarisation: applications in remote sensing. OUP Oxford.

Cloude, S.R., Pottier, E., 1997. An entropy based classification scheme for land applications of polarimetric sar. IEEE transactions on geoscience and remote sensing 35, 68–78.

Davidson, M.W., Le Toan, T., Mattia, F., Satalino, G., Manninen, T., Borgeaud, M., 2000. On the characterization of agricultural soil roughness for radar remote sensing studies. IEEE Transactions on Geoscience and Remote Sensing 38, 630–640.

De Bernardis, C.G., Vicente-Guijalba, F., Martinez-Marin, T., Lopez-Sanchez, J.M., 2015. Estimation of key dates and stages in rice crops using dual-polarization SAR time series and a particle filtering approach. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 1008–1018.

Della Vecchia, A., Ferrazzoli, P., Guerriero, L., Ninivaggi, L., Strozzi, T., Wegmuller, U., 2008. Observing and modeling multifrequency scattering of maize during the whole growth cycle. IEEE Transactions on Geoscience and Remote Sensing 46, 3709–3718.

Dey, S., Bhattacharya, A., Ratha, D., Mandal, D., Frery, A.C., 2020a. Target characterization and scattering power decomposition for full and compact polarimetric SAR data. IEEE Transactions on Geoscience and Remote Sensing .

Dey, S., Bhattacharya, A., Ratha, D., Mandal, D., McNairn, H., Lopez-Sanchez, J.M., Rao, Y., 2020b. Novel clustering schemes for full and compact polarimetric SAR data: An application for rice phenology characterization. ISPRS Journal of Photogrammetry and Remote Sensing 169, 135–151.

ESA, 2017. Sen4cap - Sentinels for common agriculture policy. `http://esa-sen4cap.org/`.

Ferrazzoli, P., Paloscia, S., Pampaloni, P., Schiavon, G., Solimini, D., Coppo, P., 1992. Sensitivity of microwave measurements to vegetation biomass and soil moisture content: A case study. IEEE Transactions on Geoscience and Remote Sensing 30, 750–756.

Fikriyah, V.N., Darvishzadeh, R., Laborte, A., Khan, N.I., Nelson, A., 2019. Discriminating transplanted and direct seeded rice using Sentinel-1 intensity data. International Journal of Applied Earth Observation and Geoinformation 76, 143–153.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: Planetary-scale geospatial analysis for everyone. Remote sensing of Environment 202, 18–27.

Hariharan, S., Mandal, D., Tirodkar, S., Kumar, V., Bhattacharya, A., Lopez-Sanchez, J.M., 2018. A novel phenology based feature subset selection technique using random forest for multitemporal PolSAR crop classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, 4244–4258.

Hird, J.N., DeLancey, E.R., McDermid, G.J., Kariyeva, J., 2017. Google earth engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. Remote sensing 9, 1315.

Homayouni, S., McNairn, H., Hosseini, M., Jiao, X., Powers, J., 2019. Quad and compact multitemporal C-band PolSAR observations for crop characterization and monitoring. International Journal of Applied Earth Observation and Geoinformation 74, 78–87.

Jackson, K.R., Ramakrishnan, L., Muriki, K., Canon, S., Cholia, S., Shalf, J., Wasserman, H.J., Wright, N.J., 2010. Performance analysis of high performance computing applications on the amazon web services cloud, in: 2010 IEEE second international conference on cloud computing technology and science, IEEE. pp. 159–168.

Jiao, X., Kovacs, J.M., Shang, J., McNairn, H., Walters, D., Ma, B., Geng, X., 2014. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. ISPRS Journal of Photogrammetry and Remote Sensing 96, 38–46.

Kumar, P., Prasad, R., Gupta, D., Mishra, V., Vishwakarma, A., Yadav, V., Bala, R., Choudhary, A., Avtar, R., 2018. Estimation of winter wheat crop growth parameters using time series Sentinel-1A SAR data. Geocarto international 33, 942–956.

Kussul, N., Lemoine, G., Gallego, F.J., Skakun, S.V., Lavreniuk, M., Shelestov, A.Y., 2016. Parcel-based crop classification in Ukraine using Landsat-8 data and Sentinel-1A data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, 2500–2508.

Lasko, K., Vadrevu, K.P., Tran, V.T., Justice, C., 2018. Mapping double and single crop paddy rice with Sentinel-1A at varying spatial scales and polarizations in Hanoi, Vietnam. IEEE journal of selected topics in applied earth observations and remote sensing 11, 498–512.

Lee, J.S., Grunes, M.R., Pottier, E., 2001. Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization sar. IEEE Transactions on Geoscience and Remote Sensing 39, 2343–2351.

Li, H., Zhang, C., Zhang, S., Atkinson, P.M., 2019. Full year crop monitoring and separability assessment with fully-polarimetric L-band UAVSAR: A case study in the Sacramento Valley, California. International Journal of Applied Earth Observation and Geoinformation 74, 45–56.

Lopez-Sanchez, J.M., Cloude, S.R., Ballester-Berman, J.D., 2012. Rice phenology monitoring by means of SAR polarimetry at X-band. IEEE Transactions on Geoscience and Remote Sensing 50, 2695–2709.

Lopez-Sanchez, J.M., Vicente-Guijalba, F., Ballester-Berman, J.D., Cloude, S.R., 2014. Polarimetric response of rice fields at C-band: Analysis and phenology retrieval. IEEE Transactions on Geoscience and Remote Sensing 52, 2977–2993.

Mandal, D., Kumar, V., Bhattacharya, A., Rao, Y.S., Siqueira, P., Bera, S., 2018. Sen4Rice: A processing chain for differentiating early and late transplanted rice using time-series Sentinel-1 SAR data with Google Earth Engine. IEEE Geoscience and Remote Sensing Letters 15, 1947–1951.

Mandal, D., Kumar, V., Lopez-Sanchez, J.M., Bhattacharya, A., McNairn, H., Rao, Y., 2020a. Crop biophysical parameter retrieval from Sentinel-1

SAR data with a multi-target inversion of Water Cloud Model. International Journal of Remote Sensing 41, 5503–5524.

Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J.M., McNairn, H., Rao, Y.S., 2020b. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. Remote Sensing of Environment 247, 111954.

Mascolo, L., Lopez-Sanchez, J.M., Vicente-Guijalba, F., Nunziata, F., Migliaccio, M., Mazzarella, G., 2016. A complete procedure for crop phenology estimation with PolSAR data based on the complex Wishart classifier. IEEE Transactions on Geoscience and Remote Sensing 54, 6505–6515.

McNairn, H., Brisco, B., 2004. The application of C-band polarimetric SAR for agriculture: A review. Canadian Journal of Remote Sensing 30, 525–542.

McNairn, H., Jiao, X., Pacheco, A., Sinha, A., Tan, W., Li, Y., 2018. Estimating canola phenology using synthetic aperture radar. Remote Sensing of Environment 219, 196–205.

McNairn, H., Tom, J., Powers, J., Bélair, J., Berg, A., Bullock, A., Colliander, A., Cosh, A., Kim, M., Ramata, S., et al., 2016. Experimental plan SMAP validation experiment 2016 in Manitoba, Canada (SMAPVEX16-MB).

Minasny, B., Shah, R.M., Che Soh, N., Arif, C., Indra Setiawan, B., et al., 2019. Automated Near-Real-Time Mapping and Monitoring of Rice Extent, Cropping Patterns, and Growth Stages in Southeast Asia Using

Sentinel-1 Time Series on a Google Earth Engine Platform. Remote Sensing 11, 1666.

Nasrallah, A., Baghdadi, N., El Hajj, M., Darwish, T., Belhouchette, H., Faour, G., Darwich, S., Mhawej, M., 2019. Sentinel-1 data for winter wheat phenology monitoring and mapping. Remote Sensing 11, 2228.

Nelson, A., Setiyono, T., Rala, A.B., Quicho, E.D., Raviz, J.V., Abonete, P.J., Maunahan, A.A., Garcia, C.A., Bhatti, H.Z.M., Villano, L.S., et al., 2014. Towards an operational SAR-based rice monitoring system in Asia: Examples from 13 demonstration sites across Asia in the RIICE project. Remote Sensing 6, 10773–10812.

Nemani, R., Votava, P., Michaelis, A., Melton, F., Milesi, C., 2011. Collaborative supercomputing for global change science. Eos, Transactions American Geophysical Union 92, 109–110.

Nguyen, D.B., Gruber, A., Wagner, W., 2016. Mapping rice extent and cropping scheme in the Mekong Delta using Sentinel-1A data. Remote Sensing Letters 7, 1209–1218.

Periasamy, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1. Remote Sensing of Environment 217, 537–549.

Ratha, D., Mandal, D., Kumar, V., McNairn, H., Bhattacharya, A., Frery, A.C., 2019. A generalized volume scattering model-based vegetation index from polarimetric SAR data. IEEE Geoscience and Remote Sensing Letters 16, 1791–1795.

Redkar, T., Guidici, T., Meister, T., 2009. Windows azure platform. Springer.

Sarabandi, K., 1991. Electromagnetic scattering from vegetation canopies. .

Schlund, M., Erasmi, S., 2020. Sentinel-1 time series data for monitoring the phenology of winter wheat. Remote Sensing of Environment 246, 111814.

Singha, M., Dong, J., Zhang, G., Xiao, X., 2019. High resolution paddy rice maps in cloud-prone Bangladesh and Northeast India using Sentinel-1 data. Scientific data 6, 1–10.

Song, Y., Wang, J., 2019. Mapping winter wheat planting area and monitoring its phenology using Sentinel-1 backscatter time series. Remote Sensing 11, 449.

Steele-Dunne, S.C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.W., Papathanassiou, K., 2017. Radar remote sensing of agricultural canopies: A review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10, 2249–2273.

Trudel, M., Charbonneau, F., Leconte, R., 2012. Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarization data for estimating soil moisture over agricultural fields. Canadian Journal of Remote Sensing 38, 514–527.

Ulaby, F., 1975. Radar response to vegetation. IEEE Transactions on Antennas and Propagation 23, 36–45.

Ulaby, F.T., El-Rayes, M.A., 1987. Microwave dielectric spectrum of vegetation-Part II: Dual-dispersion model. IEEE Transactions on Geoscience and Remote Sensing , 550–557.

Van Tricht, K., Gobin, A., Gilliams, S., Piccard, I., 2018. Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: a case study for Belgium. Remote Sensing 10, 1642.

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. Remote Sensing 10, 1396.

Wali, E., Tasumi, M., Moriyama, M., 2020. Combination of Linear Regression Lines to Understand the Response of Sentinel-1 Dual Polarization SAR Data with Crop Phenology—Case Study in Miyazaki, Japan. Remote Sensing 12, 189.

Wang, H., Magagi, R., Goïta, K., Trudel, M., McNairn, H., Powers, J., 2019. Crop phenology retrieval via polarimetric sar decomposition and random forest algorithm. Remote Sensing of Environment 231, 111234.

Whelen, T., Siqueira, P., 2018. Time-series classification of Sentinel-1 agricultural data over North Dakota. Remote sensing letters 9, 411–420.

Wiseman, G., McNairn, H., Homayouni, S., Shang, J., 2014. RADARSAT-2 polarimetric SAR response to crop biomass for agricultural production monitoring. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 4461–4471.

Wu, L.k., Moore, R.K., Zoughi, R., 1985. Sources of scattering from vegetation canopies at 10 Ghz. IEEE Transactions on Geoscience and Remote Sensing , 737–745.