Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

Using the surface scattering mechanism from dual-pol SAR data to estimate topsoil particle-sizefractions

Sandra Cristina Deodoro [a,*], Rafael de Andrade Moral [b], Réamonn Fealy [c], Tim McCarthy [d], Rowan Fealy [a]

[a] Irish Climate Analysis and Research Units (ICARUS), Department of Geography, Maynooth University, Co. Kildare, Ireland
[b] Department of Mathematics & Statistics, Maynooth University, Co. Kildare, Ireland
[c] Teagasc Agrifood Business and Spatial Analysis Department, Ashtown, Co. Dublin, Ireland
[d] National Centre for Geocomputation, Maynooth University, Co. Kildare, Ireland

ARTICLE INFO

ABSTRACT

Data extracted from Synthetic Aperture Radar (SAR) have been widely employed to estimate soil properties. However, these studies are typically constrained to bare soil conditions, as soil information retrieval in vegetated areas remains challenging. Polarimetric decomposition has emerged as a potentially useful method to separate the scattering contributions of different targets (e.g. canopy/leaves and the underlying soil), which is of significance for areas that are near-permanently covered in low-lying vegetation (e.g. grass) like Ireland – the study area for this investigation. Here, we test the surface scattering mechanism, derived from H-alpha dual-pol decomposition, together with other covariates, to estimate percentages of sand, silt, and clay, over vegetated terrain, using Sentinel 1 data (dual-pol C-band SAR). The statistical modelling approaches evaluated – linear regression (LRM) and tree-based regression models (machine learning) – explicitly consider the compositional nature of soil texture. When compared to the models fitted without surface scattering data, results showed that the inclusion of the surface scattering data improved estimates of silt and clay, with the compositional linear regression model, and estimates of sand and silt fractions with different tree-based models. While not without limitations, our study demonstrated that the polarimetric decomposition method, which is typically used for classification and segmentation purposes, could also be used for soil property estimation, broadening the application of this technique in microwave remote sensing studies.

1. Introduction

Researching and developing novel methods for mapping soil properties contributes to advances in soil science. Geotechnologies, including remote sensing techniques, play an important role in this task, for example, by exploiting methods and ancillary variables for compositional data interpolation to improve model performance (Wang and Shi, 2017; Wang and Shi, 2018; Mondejar and Tongco, 2019). In turn, the importance of using remote sensing data and techniques to analyse and estimate soil properties relies on their use over large areas, where achieving similar coverage taking field measurements are costly and time-consuming, and may be limited by access to sites. Furthermore, optical- and microwave remote sensing have significantly contributed to the advancement of soil science across a range of different applications,

such as the estimation of soil particle size-fractions (Gholizadeha et al., 2018; Bousbih et al., 2019; Domenech et al., 2020; Marzahn and Meyer, 2020; Azizi et al., 2023), soil moisture (Jackson and Schmugge, 1989; Barrett et al., 2007; Barrett et al., 2009; Pratola et al., 2014; Petropoulos et al., 2015), soil carbon content (Gholizadeha et al., 2018; Žížala et al., 2019; Katebikord et al., 2022), soil organic matter (van Wesemael et al., 2023; Yang et al., 2023), soil salinity (Metternicht and Zinck, 2003; Khajehzadeh et al., 2022) and soil mineralogical composition (Coblinski et al., 2021).

In the microwave remote sensing realm, synthetic aperture radar (SAR), such as Sentinel 1, can retrieve biophysical measures of vegetation and soil by the echoes from the Earth's surface through a radar antenna, measured as backscattering. In soils, the sensitivity of SAR relies on the interaction between the transmitted electromagnetic wave

and the soil moisture and soil roughness (Jagdhuber, 2012). In the microwave region of the electromagnetic spectrum, natural materials have a dielectric constant ranging from 3 to 8 when dry, whereas water has a value of approximately 8 (ESA, 2007:25) (Supplementary table 1). As water has a dielectric constant 10 times that of dry soil, the presence of water in the top few centimetres of unvegetated soil (bare soil) can thereby be detected in radar imagery (ESA, 2007:25). In wet soil, the dielectric constant value is ~ 20 (Barrett et al., 2007).

Studies on the importance of soil properties, including soil texture and soil PSF, on the dielectric constant date back to the 1980 s, conducted by Schmugge (1980), Wang and Schmugge (1980), Dobson et al (1985), and Jackson (1987) who developed dielectric mixture models for soil. The dielectric constant is related with the bound water in soil surfaces which in turn is dependant on the total surface area of the soil available to the water molecules. Jackson (1987) demonstrated that a sand with a specific soil surface of 1 m^2/g would have a much smaller bound water fraction than a clay with a value of 300 m^2/g, at a microwave frequency of 1.4 GHz (Supplementary Fig. 1).

The radar backscatter signal has a different relationship with the soil dielectric (moisture, and physical properties) and surface roughness. For example, Baghdadi et al. (2016) indicated that VV and VH intensities are more sensitive to surface roughness (i.e. k-root mean square height) than change in moisture, being more significant for higher incidence angles (45°) (Supplementary Fig. 2). Sentinel-1 with Interferometric Wide (IW) swath mode acquisition, which is the product employed in this study, has an incidence angle ranging from 29.1° to 46.0° (ESA, n.d.).

Soil texture is important for remote sensing (either optical or microwave types), particularly considering soil mineralogy, since clayey minerals (hematite, goethite, kaolinite, chlorite, and illite to name a few) affect soil reflectance (Coblinski et al., 2021), dielectric properties (permittivity) (Hallikainen et al., 1985; Das and Paul, 2015; Babaeian et al., 2019) and electrical conductivity (Fletcher, 2022). Clayey minerals influence the water dynamics in soil (Coblinski et al., 2021) and both chemical and physical soil properties.

Regarding wave polarisations (e.g. VV, VH), polarimetry is an important concept in SAR applications, which relies on oscillations of the transmitted/received signals. The interaction of the transmitted pulse with the surface scatterers modifies the polarisation of the transmitted signal. Such modification depends on the geometry (size, shape, orientation) and the physical properties (dielectric content, density) of the scatterers (Jagdhuber, 2012). In the context of soils, such differential propagation of microwaves causes interferometric effects that lead scattering elements on soil to produce phase and coherence variations (De Zan et al., 2014). Polarimetry thereby is a potentially useful technique for retrieving soil properties as it provides information about the physical description of scattering processes (Jagdhuber, 2012). Such a technique builds upon interferometry as it uses the phase and amplitude information of the wave. In the soil context, the differential propagation of electromagnetic waves into the soil causes interferometric effects, meaning that small scattering elements in the soil or moisture gradients can produce phase and coherence variations (De Zan et al., 2014). Due to the unique characteristic of SAR polarimetry, polarimetric decomposition is a common technique to separate scattering processes for use in segmentation and classification studies (Cloude and Pottier, 1996; Freeman and Durden, 1998; Ji and Wu, 2015; Abdel-Hamid et al., 2018; Husman et al., 2021).

To date, the investigation of soil parameters with SAR has largely been constrained to bare soil areas and for retrieving soil moisture or soil roughness (Bousbih et al., 2019; Marzahn and Meyer, 2020; Bhogapurapu et al., 2022). Notwithstanding the importance of soil texture and PSF as a proxy for other soil properties estimates and various applications (e.g. agriculture, erosion, civil engineering, etc), the use of SAR to estimate these remain underexploited.

Separating individual scattering contributions between soil and vegetation under vegetated fields is a pivotal task to extracting the soil properties (Jagdhuber, 2012). This is of significance for areas that are near permanently covered in low-lying vegetation (e.g. grass) like Ireland. In areas covered with some vegetation, the SAR backscatter comprises of scattering from soil, vegetation, and the interaction of these two elements (Dou et al., 2022). However, removing the contribution of vegetation to capture only the topsoil surface scattering is challenging (Wang et al., 2017; Marzahn and Meyer, 2020; Bhogapurapu et al., 2022). In this context, polarimetric decomposition emerges as a potential method, as it separates the scattering contributions of different targets (e.g. canopy and underlying soil) (Barrett et al., 2009; Jagdhuber, 2012; Jagdhuber et al., 2013; Ji and Wu, 2015; Wang et al., 2017; Mascolo et al., 2022). In optical remote sensing, such technique is similar to the spectral mixing analysis or spectral unmixing methods used to decompose a reflectance or corrected radiance source spectrum into a set or a given endmember spectra or spectral signature, since not all pixel of a natural scene is pure, but a mixture (Somers et al., 2016; Manolakis et al., 2016; Halbgewachs et al., 2022).

Polarimetric decomposition or target decomposition is the process of extracting information about the scattering process and describing scattering properties by partitioning the total power or intensity into relative contributions of different idealised scatterers. The H-alpha Pol Decomposition (H-α decomposition), otherwise known as Cloude-Pottier Decomposition, was developed by Cloude and Pottier (1997) to separate the scattering contributions of targets in radar images. Similar to other polarimetric decomposition methods, it converts the random scattering mechanism into several independent elements that can be related to the physical properties of the target (Salma et al., 2022; Cloude and Pottier, 1997), whose information is given by the resulting alpha, isotropy, and alpha parameters. The assumption behind this method is that there is always a dominant scattering mechanism in each pixel or cell, which is typically provided by the alpha parameter. As a target decomposition, the resulting scattering mechanisms can be surface, volume or double-bounce types.

Surface scattering refers to the scattering that occurs on the surface between two different but homogeneous media (e.g. soil surface or topsoil, sea surface) and the direction of scattering depends on the degree of surface roughness. When electromagnetic radiation transmits from one medium to another, volume scattering occurs. Trees or branches, subsurface or deeper soil layers, snow layers, etc. are examples of volume scattering. The surface and the volume components come from the ground surface and the vegetation, respectively (Magagi et al., 2022). In general, single scatterers will have lower alpha values being produced by a rough surface (i.e. seawater, snow surface, ploughed bare soil) which will increase as biomass increases. On a bare, smooth field, alpha will indicate a dominance of single bounce scattering and entropy will remain low. With increasing vegetation cover (e.g. dense vegetation, forests), the polarisation of the scattered wave becomes less predictable and entropy will increase.

While H-alpha Decomposition was originally developed for quad-polarised SAR, the method has been adapted for dual-polarised radar, such as Sentinel 1. Despite inherent limitations due to the reduced number of polarisations (i.e. two channels), dual-polarisation has produced reasonable results in different applications as outlined in the next section. Here, we test the potential importance of surface scattering, derived from a dual-polarimetric decomposition method (H-alpha decomposition), to estimate the content (%) of sand, silt, and clay from in situ topsoil samples from bare-soil and low vegetated cover employing a linear regression and a tree-based model with additional covariates. An intrinsic characteristic of soil particle size-fractions (PSF) – i.e. sand, silt, and clay – is their compositional nature (i.e. the sum of the fractions-components is equal to 100). This requires appropriate multivariate analysis that can handle compositional data in order to avoid the potential for spurious correlation or unrealistic results (Aitchison, 1982; Aitchison, 2005; Niang et al., 2014; Pawlowsky-Glahn et al., 2015; Filzmoser et al., 2018; Todorov, 2021). Moreover, the sum of the components in a predicted soil sample (e.g. 100 %) is not guaranteed with standard statistical methods that are designed for

unconstrained multivariate data in a Euclidian space (Weiss et al., 2017; Amirian-Chakan et al., 2019). We explicitly consider the compositional nature of soil texture in the modelling approach.

We sought to address two key research questions: (i) to evaluate to what extent polarimetric decomposition with scattering partitioning can be employed to predict soil particle-size fractions, that is, to assess how well the Sentinel 1 pseudo-polarimetric product contributes to the estimates of soil PSF in comparison with the original dual-polarization product SLC-1; and, (ii) to evaluate if the surface scattering mechanism can ameliorate vegetation cover issues for soil particle size prediction on soils under (low) vegetated cover.

It is worth noting that, due to the limited polarisation channels, this work does not intend to be exhaustive in examining the efficiency of the method to accurately distinguish targets, rather it seeks to explore and expand this method for soil PSF prediction according to the research questions. Thus, limitations and future works are highlighted based on our findings.

This paper is organised into the following sections. Section 2 briefly presents the theoretical background and some related works. Section 3 outlines the methodological procedure and datasets employed, including a description of the study area. Sections 4 and 5 present the results and discussion, respectively. Finally, Section 6 draws some conclusions from the work.

2. Theoretical background and related works

2.1. Radiometric calibration of SAR image

Prior to employing SAR data, the radar data must be calibrated, since this procedure provides a measure of radar reflectivity of a surface (i.e. the radar backscatter). Radiometric calibration corrects a SAR image so that the pixel values represent backscattering from the Earth surface, being a crucial procedure for a quantitative use of SAR data.

The radiometric corrections are purposes-specific. For interferometric and polarimetric analyses – the later includes polarimetric decomposition – the output data need to be in complex number format (i.e. real and imaginary bands) to obtain coherence and covariance matrices. For other applications, radiometric calibration converts backscatter intensity as received by the sensor to the normalized radar cross section (sigma nought or σ^0).

Backscattering coefficients are reliant on the local incidence angle and local topography and related to characteristics of the targets, such as physical (geometric and electromagnetic), soil moisture, surface roughness, and dielectric constant (Imperatore and Di Martino, 2023). Abrupt land use land cover changes (e.g. from grass or bare soil to trees) may also affect backscattering coefficients when the backscatter and incidence angle behaviour varies over the image (Freeman, n.d.). The radiometric corrections are thereby also area-specific.

There are three different calibrations required for Sentinel-1 data, depending on the plane reference area used to normalise the backscatter (Fig. 1). Sigma nought (σ^0) refers to the backscatter returned to the antenna from a unit area on the ground, being related to ground range (i. e. normalised to ground area) and based on ellipsoid (or incidence angle) corrected SAR backscatter. It is useful for flatter terrain and is commonly used to analyse surface scattering and surface properties (Barrett et al., 2007; Zhang et al., 2017; Periasamy, 2018; Nasirzadehdizaji et al., 2019; Mandal et al., 2020; Qu et al., 2020; Salma et al., 2022). Beta nought (β^0) is normalised to a slant range plane area and refers to the radar brightness coefficient, which is the ratio between the power transmitted and received by the antenna. Gamma nought (γ^0) or RTC Gamma nought (Small, 2011) is normalised to an area perpendicular to the line of sight and accounts for terrain variations, advantageous over undulated terrain, mountainous and hilly areas. An recent attempt to improve Sentinel-1 Radiometrically Terrain Corrected (RTC) to reduce processing and storage through orbital stability is outlined in Navacchi et al. (2023).
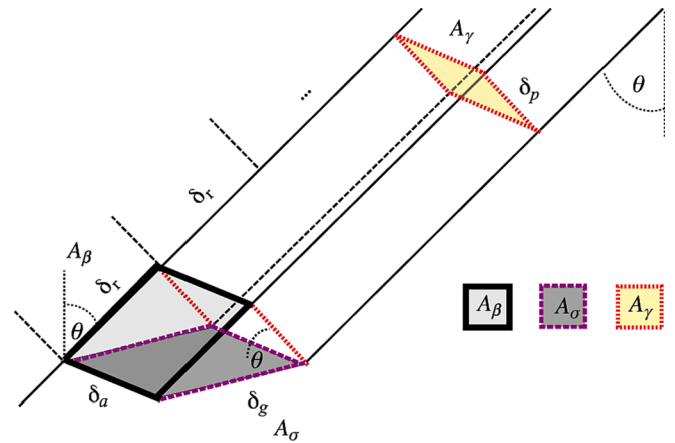

Fig. 1. Normalization areas for SAR backscatter (Small, 2011). β0 refers to reflectivity in an area in slant range and measures brightness in SAR images; σ^0 is illumination corrected and measures the power returned to antenna from ground; γ^0 is plane perpendicular to the local look direction and is radiometrically corrected being suitable for volume scatterers.

Thus, for a spaceborne SAR, the σ^0 and the γ^0 backscattering coefficients relies on a flattened earth model and a curved earth model, for the receive and transmit polarizations of the radar (Freeman, n.d).

2.2. The H-alpha dual pol decomposition

Polarimetric decomposition is the process of partitioning the total intensity into contributions of different scatterers to extract information about the scattering properties of targets. The H-alpha Pol Decomposition (H-α decomposition), otherwise known as Cloude-Pottier Decomposition, was developed by Cloude and Pottier (1997). Similar to other polarimetric decomposition methods, it converts the random scattering mechanism into several independent elements that can be related to the physical properties of the target (Cloude and Pottier, 1997; Salma et al., 2022). One advantage of this method is the lack of physical constraints imposed by assumptions of a particular statistical distribution (Cloude and Pottier, 1997). The assumption behind this method is that there is always a dominant scattering mechanism in each pixel or cell. Based on this, Cloude and Pottier (1997) employed a three-symbol Bernoulli model to obtain estimates of the average parameters for the dominant target using POLSAR data (AIRSAR) from NASA's Jet Propulsion Laboratory (JPL).

Typically, there are five different target polarimetric statistical-based descriptors: Sinclair Matrix [S], Kennaugh Matrix [K], Target Vectors ($\underline{k}$, $\underline{\Omega}$), Coherence Matrix [T], and Covariance Matrix [C]. The H-alpha Decomposition pertains to the coherence matrix group (Lee and Pottier, 2009), whose coherence matrix allows analysing single scattering mechanisms (i.e. surface scattering) within a pixel as well as their contribution to the total signal. The H-alpha Pol Decomposition method relies on the eigen decomposition of the coherency matrix [T3] into types of scattering processes (the eigenvectors) and their relative magnitudes (the eigenvalues) (Lee and Pottier, 2009). It is an eigenvectors-eigenvalues-based analysis (Equations 1–4) used to examine the intensity (coherency matrix) of the scattering matrix.

The method produces three secondary parameters – Entropy (H) (Eq.1), Anisotropy (A) (Eq. (2) and Alpha angle (α) (Eq. (3), which contain information on the scattering, from the eigen decomposition of the coherency matrix (Lee and Pottier, 2009).

$$H = \sum_{i=1}^{3} -P_i log_3(P_i), P_i = \lambda_i \bigg/ \sum \lambda_k \qquad (1)$$

$$A = (\lambda_2 - \lambda_3)/(\lambda_2 + \lambda_3) \qquad (2)$$

$$\alpha = -\sum_{i=1}^{3} P_i \alpha_i \qquad (3)$$

where λ_1, λ_2, and λ_3 be the eigenvalues of the coherency matrix ($\lambda_1 > \lambda_2 > \lambda_3 > 0$). From this, the corresponding eigenvectors [$\underline{u_1}$, $\underline{u_2}$, $\underline{u_3}$]^{*T} are expressed as follows:

$$u_i = \left[\cos\alpha_i, \sin\alpha_i \cos\beta_i e^{i\delta}, \sin\alpha i \sin\beta_i e^{i\gamma} \right]^T \qquad (4)$$

where α_i denotes the scattering mechanisms of a target; β_i is the orientation angles; δ, and γ represents the phases.

Since the method was adapted from a quad-pol SAR equation, the parameters and indices in equations 1–4 that refer to four polarisations are reduced to 2 to be suitable for use with two polarisations in a coherence matrix method, since both the scattering matrix and scattering vector of VV-VH or HH-HV SAR, for example, has two elements.

Entropy indicates how random the scattering mechanisms are and lies between 0 and 1 where H = 0 exhibits a pure target (single scattering element) and polarised surface (Harfenmeister et al., 2021). If entropy is greater than zero, thereby, there is an indication of the presence of two different mechanisms, meaning that the scattering mechanisms are not surface, but a mixture of two different mechanisms. Anisotropy provides information on the relative effect of the second and third mechanisms and ranges between 0 and 1, being therefore complementary to entropy (Harfenmeister et al., 2021), especially when H > 0.7 in full-pol SAR, since H < 0.7 the anisotropy is noisy and has limited importance (dos Santos et al., 2009). The Alpha parameter is used to determine the dominant scattering mechanism and varies between 0° and 90°, particularly in full polarisation SAR. It is the main parameter (Cloude and Pottier, 1997) as it provides information related to the surface, volume (e.g. trees) and double-bounce scattering (e.g. urban areas).

Complementary to the polarimetric decomposition, there is the H-alpha plane plot that builds upon the *H-α* variation for a coherency matrix and consists of nine zones representing different scattering behaviour (Cloude and Pottier, 1997) (Supplementary Fig. 3). These zones result from the interplay between Alpha and Entropy. An example of this polarimetric decomposition method applied to an agricultural field is depicted in Supplementary Fig. 4a. It is worth noting that the H-alpha plane for dual-pol differ from full-pol in terms of the boundary lines and zone thresholds (Supplementary Fig. 4b) as demonstrated by Ji and Wu (2015). A summary of the physical scattering characteristics of each of the nine zones in the H-α classification plane is found in Supplementary Table 2.

Regarding applications of the method, Jagdhuber et al. (2013) underline the success of polarimetric decomposition to separate the vegetation contribution and estimate soil moisture on isolated bare ground components, enabling the interpretation and decomposition of different scattering contributions. The method has also been employed for bio-parameter extraction such as biomass (Suman, 2022), rice monitoring (Koppe et al., 2013) and soil moisture (Magagi et al., 2022). Qu et al. (2022) demonstrated the use of polarimetric decomposition from quad-pol SAR data for land cover classification.

In the context of dual-pol SAR, some limitations of the method have been reported in the literature. For example, the closer alpha is to 0°, the greater the chance of surface scattering. However, some studies demonstrated that it is not possible to distinguish pure surface scattering (i.e. alpha angle = 0) from dihedral or double-bounce scattering (alpha = 90) with S1 data due to reduced polarisation or lack of the remaining quad-pol (e.g. Engelbrecht et al., 2017; Mascolo et al. 2022; Salma et al., 2022); that is, alpha values for the ground may not be as low as 0.0. According to the dual-pol covariance matrices of the surface and dihedral targets, (i.e., C_surface = [1, 0; 0, 0] = C_double-bounce,) it is not possible to distinguish them. As a result, the scattering mechanisms can not be separated. Cross-polar channel (VH or HV), which is sensitive to vegetation, becomes a poor estimate of volume scattering (Mascolo

et al., 2022), particularly under vertical dipole situations, hindering the separation of soil and vegetation. While the alpha values vary from 0° to 90° in quad-polarization, Salma et al. (2022) obtained a maximum alpha value of approximately 45°, in a study area situated in India, with Sentinel-1 dual-pol. In our study area located in Ireland (Fig. 2), alpha values mostly ranged from 0.051 to approximately 39, for the whole area – Supplementary Fig. 5.

Despite inherent limitations due to the reduced number of polarisations, dual-polarisation has been employed in a number of different applications. For example, Ji and Wu (2015) analysed the abilities of dual-polarisations (HH-VV, HH-HV, and HV-VV) from different SAR sensors to discriminate scattering mechanisms, namely surface, volume and double bounce scattering. Despite the lack of cross-polarization, they demonstrated that dual-pol HH-VV could effectively derive the scattering mechanisms; however, the HH-HV and HV-VV could only partially extract low, medium, and high entropy scattering mechanisms due to the lack of co-polarization. Their work builds upon theoretical and experimental results comparing performance with that of fully polarimetric SAR.

Mascolo et al. (2022) applied dual-pol decomposition to Sentinel-1 data over different land cover types, including agricultural, forest, urban, and glacial land-ice. They showed that some elements of physical modelling can be used to better exploit information from Sentinel 1 and that the dual-pol decompositions can be used to extract important underlying physical information, provided that the correct interpretation of parameters is used. Notwithstanding the benefits of this method, they also considered the limitations associated with the use of dual polarisation SAR (e.g. the selective availability of different combinations of VH-VV and HH-HV and their impact on the interpretation of the decomposition outputs and identification of phenomena in an image). These authors also demonstrated the importance of the VH and VV backscattered signal for separating the land-cover types and for the estimation of important crop variables, based on biomass and phenology.

Salma et al. (2022) analysed target properties for a range of crop types, including ginger, tobacco, rice, cabbage, and pumpkin crops, using H-A-α dual-polarization decomposition with a temporal series of Sentinel 1 images (spanning 19th January and 27th September 2020). They found variation from the quad-polarisation decomposition, where the alpha values fell between 0° and 90°, whereas the alpha values of dual-polarisation data were found to vary between 7.92° and 45.23°. They also observed that the corresponding pixels on the ground exhibited surface scattering over the H-α plane. Finally, they found that the crop growth scattering mechanism on the H-α plane produced similar results to the temporal analysis.

The high frequency overpass of Sentinel 1 together with the sensitivity of polarimetric parameters (entropy, anisotropy, and alpha) to biophysical properties have increased the use and application of S1 for monitoring biophysical parameters over agricultural fields, as demonstrated by Dave et al. (2023) and Harfenmeister et al. (2021), and for crop classification (e.g. Husman et al., 2021; Woźniak et al., 2022). An attempt to improve interference between different scattering components in distributed scatterer employing Sentinel 1 is found in Zhang et al. (2022).

3. Data and methods

3.1. Study region

The study area encompasses a central swath of the Republic of Ireland (Fig. 2). Agriculture is the dominant land use land cover in Ireland, accounting for 67 % of the land area, from which pasture is the main agricultural class (55 %) (CORINE, 2018). Grass pasture is typically grazed by cattle on a rotation basis and/or harvested 2–3 times per year to provide a source of overwinter feed. Both management practices maintain a minimum grass height of 4 cm.
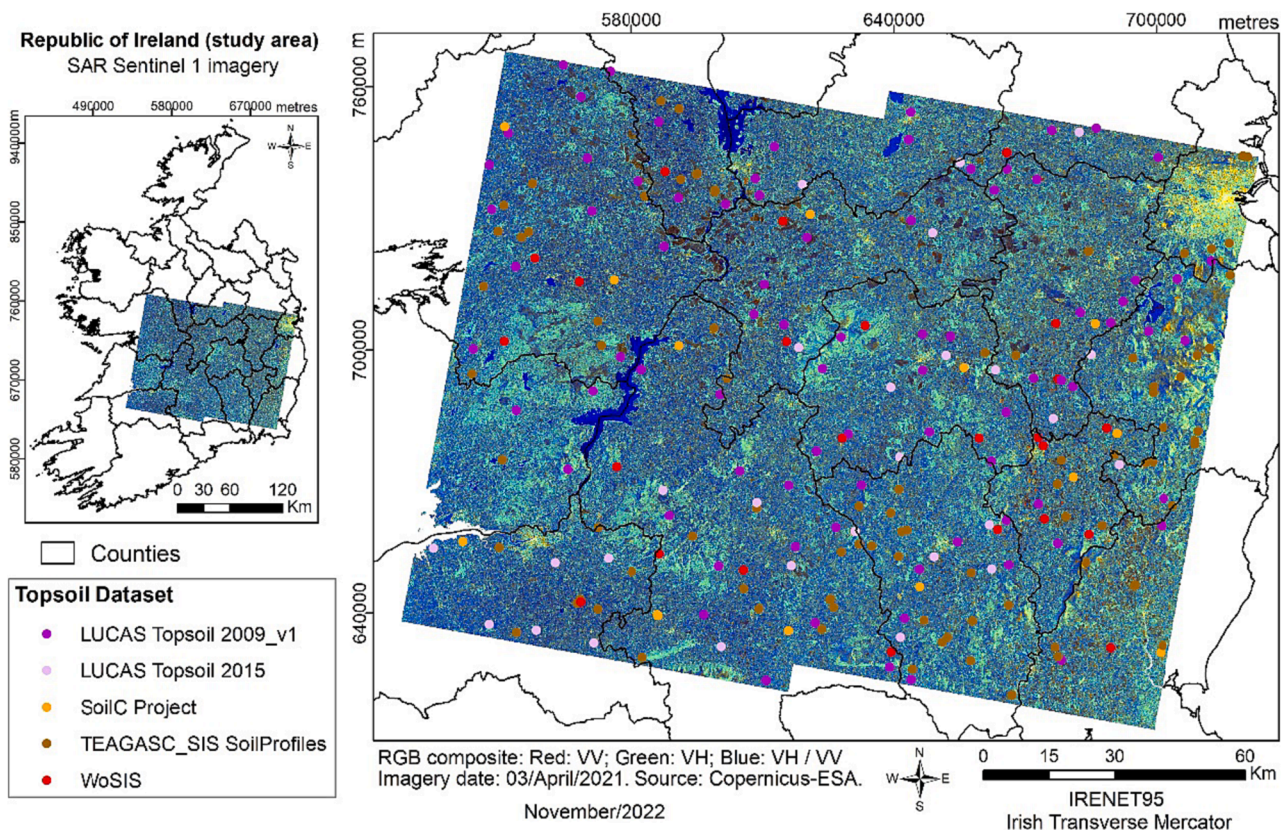
Fig. 2. The land use-land cover categories from an RGB composite applied to the Sentinel 1 imagery over the study area. Topsoil sampled sites in the study area (points) overlaid on RGB composite using VV and VH polarisations from Sentinel 1-C band. The RGB resulting image highlights some of the land use/land cover categories, such as man-made structures and urban areas (yellow) within which Dublin city is highlighted on the top-right of the figure, water bodies (blue), and bogs/peatlands/bare soil (brown), dense and higher vegetation are displayed in light cyan, and grassland, crops and low vegetation cover are shown in medium and dark shades of cyan. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Topographically, the study area consist of uplands occurring in the middle-east of the domain with geoforms oriented at W-N, N-E and S-W, whereas the lowlands are mostly present in the mid-west with W-N, E-S and N-E faced geoforms (Fig. 3). In terms of slope, mountainous areas are found in the northern and western areas, while flatter areas/lower relief are observed in the northern and southeast part of the domain. It is worth noting that the soil samples selected in the modelling stage of this work are mostly located in flatter terrain and lowland areas in the north, south and southeast areas of the study region.

The geomorphological setting affects the rainfall amounts which are typically higher on the west coast. This is due to the predominant maritime airflow off the North Atlantic Ocean that interacts with orography along that coast (Fealy et al., 2011). While Ireland experiences year round rainfall, the drier seasons are spring and summer while winter and autumn are the wettest, with approximately 260 mm and 350 mm of rain, respectively (Walsh, 2012). The predominant soil texture class in Ireland is Fine Loamy (Creamer et al., 2016).

3.2. Data

3.2.1. Soil information

A dataset comprising 235 soil samples from across Ireland and collected by various institutions – Irish Soil Information System, LUCAS Topsoil 2009–2015, Soil Carbon (Kiely, 2009), and WoSIS Soil Profile Database in both shallow and deep soil layers, was used to train and validate the models. Only the topmost layer or topsoil (0–15 cm in depth) was selected for use in this study in order to be consistent with the pulsed signals of Sentinel 1. This rationale was based on the relatively lower capability of C-band radar to penetrate into the soil –

estimated to be approximately the first 4 cm of the soil layer (Babaeian et al., 2019) (Supplementary Table 3) – compared to lower frequency-based sensors, such as L-band SAR. Our approach is consistent with Read et al. (2018), who dealt with different legacy databases to predict sand (%) and clay (%) in upper soil profiles from airborne geophysical data. In their work, they constrained soil samples collected from soil profiles up to 30 cm depth for training the model and restricted them to 0–5 cm (topsoil) in external model validation, since gamma rays emanate from topsoil layer (Read et al., 2018).

3.2.2. Microwave remote sensing data

Sentinel-1 SAR (C-band, dual polarisation VV + VH, at a centre frequency of 5.405 GHz) was employed in this work as the radar-based data. It was obtained with Interferometric Wide (IW) swath mode acquisition, from the European Space Agency (ESA). Since polarimetric decomposition requires the phase information of the wave, the Single Look Complex (SLC) product was selected. The acquisition date was the 3rd April 2021. This date was selected due to the antecedent conditions (29/03/2021–03/04/2021) – which had no associated rainfall, based on information the Irish national meteorological service, Met Éireann. Moreover, the rainfall recorded for the preceding month of March, was below the long-term average and a number of meteorological stations within the study domain recorded their driest March since 2012 (Met Éireann, 2021).

The pre-processing procedures applied to the Sentinel 1 imagery – included co-registration (S1-TOPS split as subset; orbit file applying), radiometric calibration to the complex number (i.e. a measure of radar reflectivity of a surface), S1-TOPS Deburst, S1-TOPS Merge, polarimetric speckle filtering using a Boxcar method with a 7 × 7 window size, and
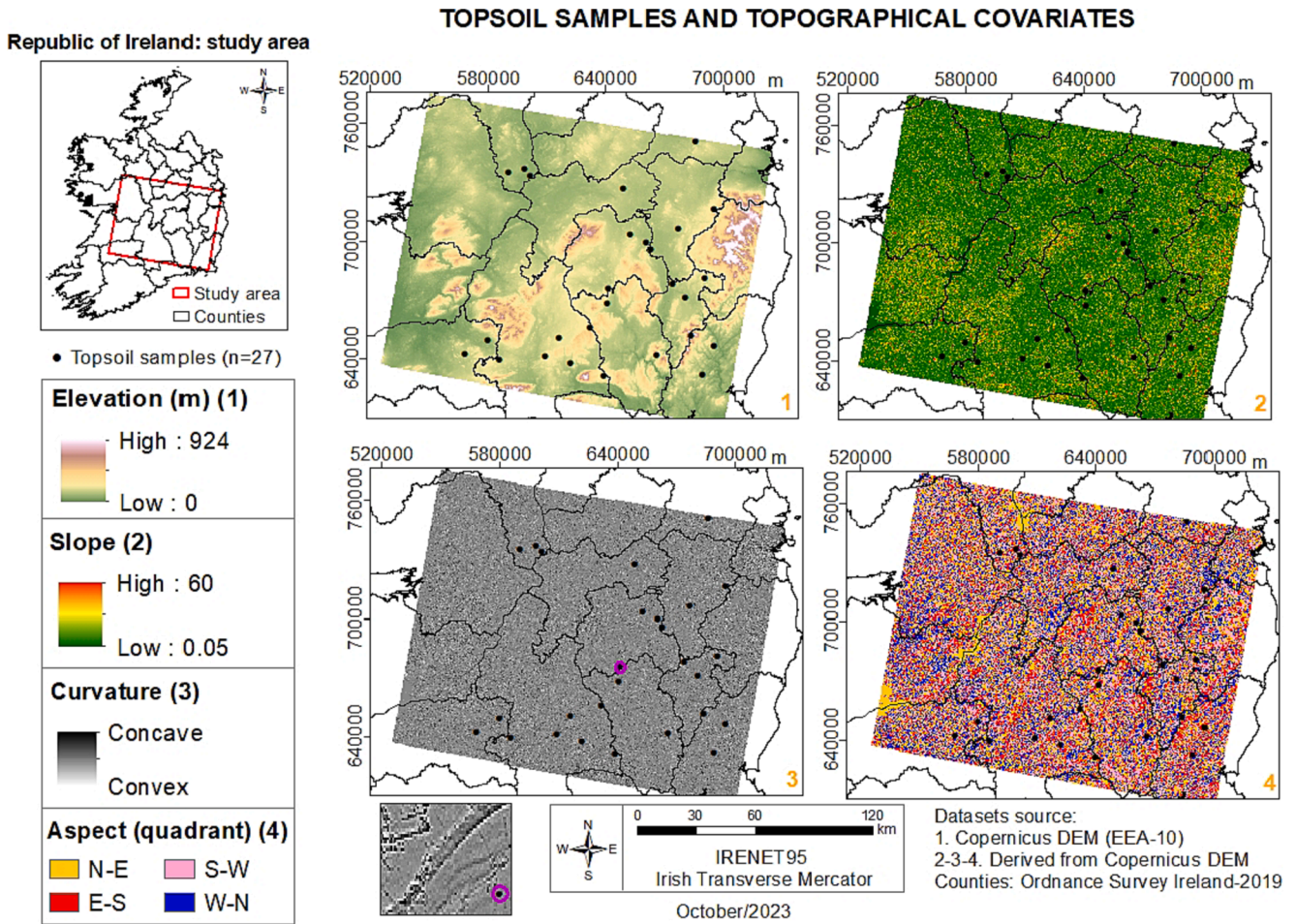
Fig. 3. Toposoil samples superimposed over topographical (morphometric) covariates images of the study area. The soil samples (n = 27) were employed in the models. Elevation data refers to the DEM-EEA-10 (10 m spatial resolution) from which the other covariates were derivate.

range-doppler terrain correction – were carried out using the Sentinel-1 Toolbox (SNAP-ESA). The complex output was selected for polarimetric processing. For calculation of the radar vegetation index, the backscattering coefficients were converted to sigma nought using the linear scale units. The Copernicus Digital Elevation Model for Europe (ESA EEA-10) at 10 m spatial resolution was used for terrain correction. All datasets were geo-coded to the Irish Transverse Mercator projection.

The radar-based data consisted of (i) VV and VH backscatter coefficients; (ii) the dual-pol radar vegetation index, (Eq. (5) (Nasirzadehdizaji et al., 2019; Gururaj et al., 2019), and (iii) topsoil surface soil moisture (SSM) measured with Sentinel-1 as part of the Copernicus programme. All data was acquired for the same date. Typically, radar vegetation indexes are useful for measuring random scattering in radar signals and values close to 0 indicate smooth, bare surfaces (Nasirzadehdizaji et al., 2019). The index employed here is an adaptation of the quad-pol RVI to dual-pol SAR and refers to the contribution of volume scattering given by cross-polarized response (Nasirzadehdizaji et al., 2019). It was employed to filter noise from the data and to reduce the vegetation effect.

$$\text{RVI} = \frac{4\sigma_{VH}^0}{\sigma_{VV}^0 + \sigma_{VH}^0} \tag{5}$$

Sentinel 2 imagery (S2), a product of optical remote sensing, was also obtained for the same date in order to compare the surface scattering from the polarimetric decomposition with soil through an RGB composite (R4G3B2 and R11G3B2).

In order to estimate the effect of soil organic matter (SOM) on the backscatter, we performed a correlation test with SOC and the SAR

backscatter. Results are presented in the Supplementary Fig. 6 and Supplementary Table 4. From the soil legacy datasets, we used SOC data from the WoSIS Database, since it contains more measurements of this soil property.

3.2.3. Topographical data

Topographical data, derived from the digital elevation model ESA EEA-10, was used to represent the soil-landscape relationship and an earth-surface characteristic related to soil-surface roughness (micro- and macro-relief of a soil surface), which is a result of variations in soil grains, soil aggregates, soil clods and tillage. In this study, topographic covariates consisted of altitude, slope, aspect and curvature (concave and convex surface), whose values where extracted from the corresponding pixels to the soil samples used in the modelling stage for predicting soil PSF. In the study area, values of altitude and slope for these points, ranged from 42 to 364 m and 0.16–11 degrees, respectively. For these sampling sites, slopes exhibit low-medium inclinations due to convex curvature and flatten summit, which justify the use of the σ^0 backscatter coefficient.

3.3. Methods

3.3.1. Dual polarimetric decomposition

The covariance 2×2 matrix (C2) was generated from the Sentinel 1. The SLC product (Eq. (6) was converted from the complex output to its covariance matrix representations (Eq. (8), to perform the polarimetric decomposition. The matrix C2 (Eq.8) consists of two real parts (C11, C22), which contains the variances, and two complex parts (C12, C21)

that contains the covariance (Salma et al., 2022).

The H-alpha Dual Pol Decomposition was further performed using the Sentinel-1 Toolbox available on the Sentinel application Platform (SNAP) in order to obtain the alpha, entropy and anisotropy values. Alpha, the main parameter (i.e. dominant scattering), was selected as a covariate in the statistical modelling as the other two parameters are highly correlated to alpha, especially entropy, with a Pearson's correlation coefficient (r) of 0.95. Thus, they were not considered in order to avoid multi-collinearity.

$$S = \begin{bmatrix} S_{VV} & S_{VH} \\ 0 & 0 \end{bmatrix} \tag{6}$$

$$k = \begin{bmatrix} S_{VV} & 2S_{VH} \end{bmatrix}^T \tag{7}$$

$$C_2 = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tag{8}$$

where S is the 2×2 complex scattering matrix or backscattering matrix (S_{VH} and S_{VV} are backscattering coefficients) (Qu et al., 2020); k is the corresponding scattering vector representation wherein T refers to the transpose of matrix; and C is the dual covariance matrix for S1-SLC (Qu et al., 2020).

In the (dual) polarisation matrix, the complex numbers C_{12} and C_{21} are conjugate, while C_{11} and C_{22} represent the energy of VV polarisation and VH polarisation, respectively. Thus, these complex numbers are displayed in an S1-SLC (Single Look Complex) file as a matrix.

Once the H-alpha Dual-Pol Decomposition was obtained, we carried out a visual inspection by overlaying, the alpha image and the Sentinel 2 image with the aid of RGB composite to identify features such as peat/bogs, water bodies, dense vegetation (forests), bare soil, ploughed soil, and soil with some degree of vegetation cover. Since we sought a distinction between soil and vegetation, only these features are indicated in the analysis.

After the visual inspection, the in situ soil sample locations were overlain on the resulting maps from the H-alpha Dual Pol Decomposition (alpha, entropy and anisotropy images). Those locations with low alpha or surface scattering were selected (blue patches in the alpha image) to extract the pixel values, following Harfenmeister et al. (2021), Mascolo et al. (2022) and Salma et al. (2022). Typically, in agricultural areas, alpha = 0° indicates a single scattering event (surface) from smooth soil or a large leaf; alpha = 45° indicates more than two scattering events from within a crop canopy (volume).

In our study area, the threshold for alpha and entropy were < 18° (low) and < 0.6 (medium), respectively (Mascolo et al., 2022). The closer alpha is to 0°, the more the surface scattering occurrence. However, there is limitation issues in distinguishing pure surface scattering from dihedral or double-bounce scattering with S1, resulting in alpha values > 0.0 for bare ground. In light of such limitations, the aforementioned threshold was set based on Mascolo et al., (2022) considering a range values of $0° < \alpha_{VHVV} < 22.5°$ and horizontal dipole (Mascolo et al., 2022). From this, the surface scattering contribution is expected due to either the spacing between plants or direct scattering from the large leaves (Mascolo et al., 2022). Surface scattering can come from vegetation, which was confirmed further from the analysis of the H-alpha plots in our study. Thus, to account for the gaps between plants in some agricultural fields, the threshold were increased from zero to include backscattering from low- and sparse- vegetation. Due to the agricultural field setting, it is difficult to have pure targets (low entropy) values that clearly distinguish them from their surroundings. Thus, they are more distributed targets and fall into the medium range for entropy. To consider the vegetation effect, that is, to reduce the effect of it on the soil backscattering, we employed the dual-pol RVI, which relies on sigma naught backscattering in its formula, in the model.

With this procedure, 27 points were ultimately selected, consisting of bare soil and low or sparse vegetation. Spectral signatures from Sentinel 2 images (obtained from 03/04/2021) were extracted for selected field locations in the study area to compare the behaviour of some features (e. g. soil, vegetation) with the decomposed targets by the H-alpha dual-pol decomposition. This is based on the fact that when there are small contributions from soil backscattering, the vegetation signature is highest in volume scattering. (Guerriero et al., 2013; Salma et al., 2022).

3.3.2. Statistical modelling

Tree-based models are alternatives to linear regression models due to their advantages of being relatively robust to overfitting and not requiring standardisation or normalisation of data; they are insensitive to the ranges in the predictor values (Zhang and Shi, 2019). Decision trees are used in classification and regression problems whose predictor space is divided up by recursive partitioning. A widely used extension to decision trees is the random forests (RF) algorithm (Breiman, 2001). Some examples of RF applications to estimate soil properties – without considering soil texture as compositional data – in a remote sensing context are found in Mirzaeitalarposhti et al. (2022), Domenech et al. (2020), Dotto et al. (2020), Cisty et al., 2019, Bousbih et al. (2019) and Ballabio et al. (2016).

Compositional data analysis (CoDA) is mostly performed in terms of log-transformations and relies on log-ratios between the parts or components of one sample. Some of the benefits of applying the log-ratio (linear) transformations are the possibility of using standard unconstrained multivariate analysis on the transformed data. Moreover, the sum of the components equal to 100 (or 1) is guaranteed when the predicted values are back-transformed to the original units (Comas-Cufí and Thió-Henestrosa, 2011). The theoretical foundations are found in Aitchison, 1982. As highlighted by Morais and Thomas-Agnan, 2021, in simplicial regression (i.e. compositional linear model; Y-LRM) one can use transformations to transport the simplex space S^D into the Euclidean space R^{D-1} to eliminate the simplex constraints problem.

Considering the compositional nature of soil texture (a composition of particle-size fractions), log-ratio transformation methods rooted in symmetry – centred log-ratio transformation (CLR), and isometric log-ratio transformation (ILR) – were employed in this study because they preserve distances. The ILR transformation was applied to the response variables or Y (sand, silt, and clay) for fitting the compositional Linear Regression Model (Y-LRM) following the default partition method in the CoDaPack software (Comas-Cufí and Thió-Henestrosa, 2011). Due to the sampling size, the model was validated by using the leave-one-out cross validation (LOOCV) method.

The CLR transformation was applied to the response variables to train a tree-based model for different extensions of tree-based algorithms, including, the traditional Random Forests-RF (Breiman, 2001), Local Linear Forests-LLF (Friedberg et al., 2021), Bayesian Additive Regression Trees-BART (Chipman et al., 2010), and SoftBART-SBART (Linero and Yang, 2018). CLR transformation was applied to the response prior to model training, since the machine learning modelling does not rely on distributional assumptions. Moreover, the singular covariance matrix resulting from the CLR-transformed data set is not an issue for the Random Forest model. We used the default method built into the CoDaPack software (Comas-Cufí and Thió-Henestrosa, 2011). For comparison with the CoDA approach, a non-CoDA-based regression was carried out for these different tree-based algorithms.

The compositional data approach has increasingly been employed to estimate soil PSF in both non-spatial (Loosvelt et al., 2013; Chappell et al., 2019; Deodoro et al., 2023) and explicitly spatial models (; Zhang and Shi, 2019; Wang and Shi, 2018; Wang and Shi, 2017; Odeh et al., 2003). Here, we employed CoDA in order to satisfy the constraint where the sum of the parts is 100. We employed two different statistical modelling approaches to see the results behaviour with a assumption-based model (LRM) and non-assumption or black-box model (tree-based model).

It is worth noting that, in addition to soil PSF estimates, the corresponding soil texture classes were also obtained. However, they were

inferred from the compositional predictions of sand, silt, and clay content rather than directly using a categorical modelling approach. The resultant soil textural classes will also be outlined in the Results section.

3.3.3. Evaluation metrics for soil particle size predictions

The performance of the modelling approaches were evaluated using three statistical indicators, as follows: Root Mean Square Error (RMSE) (Equation (9), MAE (Mean Absolute Error) (Equation (10) and Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) (Equation (11). The latter is a normalised metric that indicates the magnitude of the statistical variances between the residuals and the measured data (Nash and Sutcliffe, 1970). It measures the agreement between observed versus predicted (simulated) data in relation to the 1:1 line according to the following: NSE = 1 indicates a great fit of the model to the observed data; NSE = 0 indicates that the accuracy of predictions is given by the mean of the observed data; - Inf < NSE < 0 shows that the observed mean is predicting better than the model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \tag{9}$$

where y_i is the observed value of the response variable, $\widehat{y}_i$ is the estimated value, and n is the number of observations.

$$MAE = 1ni = 1nyi - yi \tag{10}$$

where y_i, $\widehat{y}_i$, and n represents the same descriptors that those of RMSE.

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - \overline{Y^{obs}}\right)^2}\right] \tag{11}$$

where Y_i^{obs} is the ith observation, $Yisim$ is the ith simulated (predicted) value, $\overline{Y_i^{obs}}$ is the mean of observed data, and n is the total number of observations.

4. Results

The calculated Alpha values (Fig. 4) indicate that the study area is mostly represented by volume scattering mechanisms (yellow patches) indicating occurrence of vegetation, consistent with the land cover conditions (Fig. 5). For the entire study area, Alpha values varied from
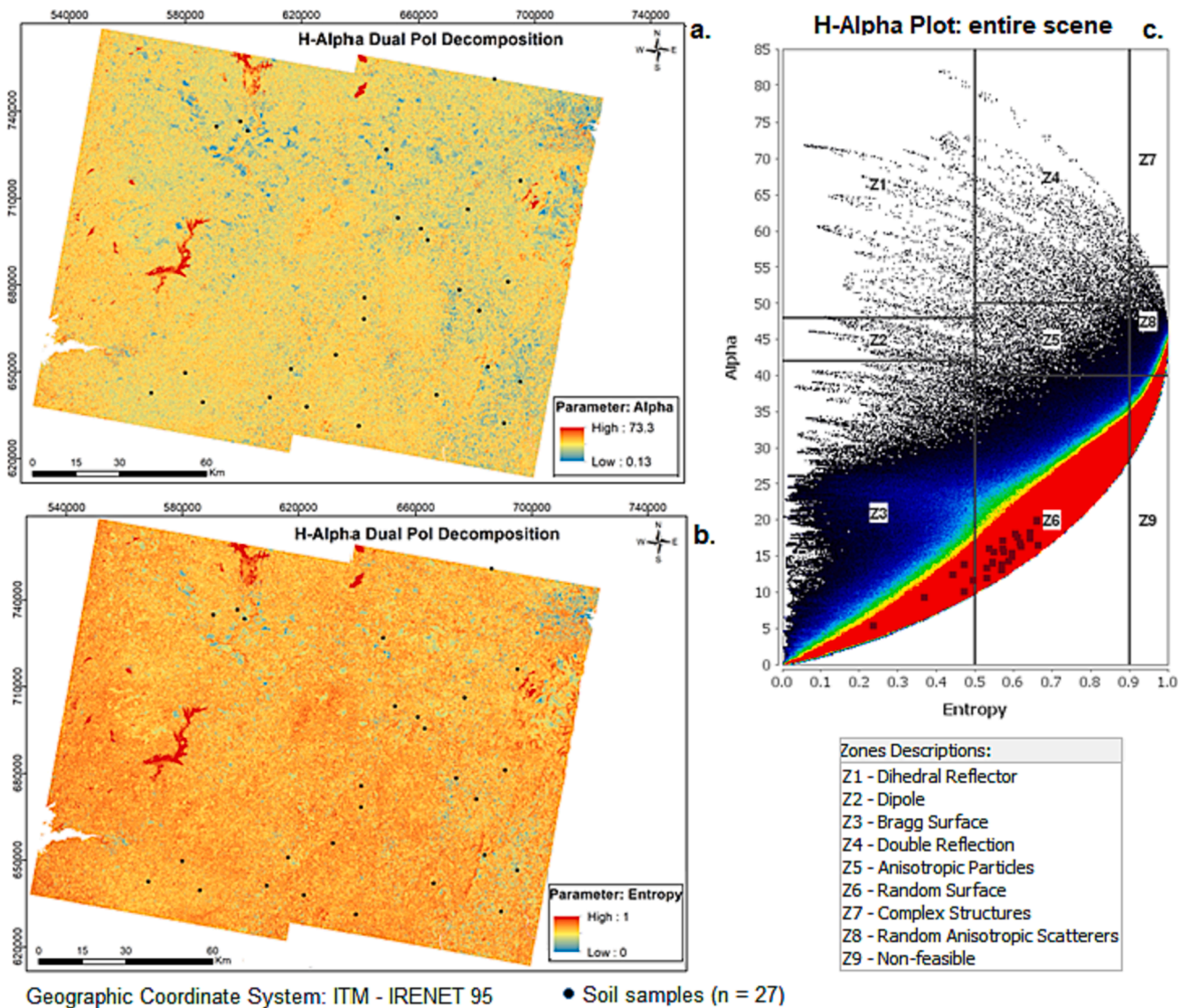

Fig. 4. Alpha and entropy maps of the entire scene (figures a, b) as retrieved from Sentinel 1 data using the dual H-alpha pol-decomposition. The corresponding H-Alpha plot for the entire scene (figure c). Square dots represent the soil samples selected (n = 27). Blue: surface scattering; yellow: volume scattering; red: double bounce scattering. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
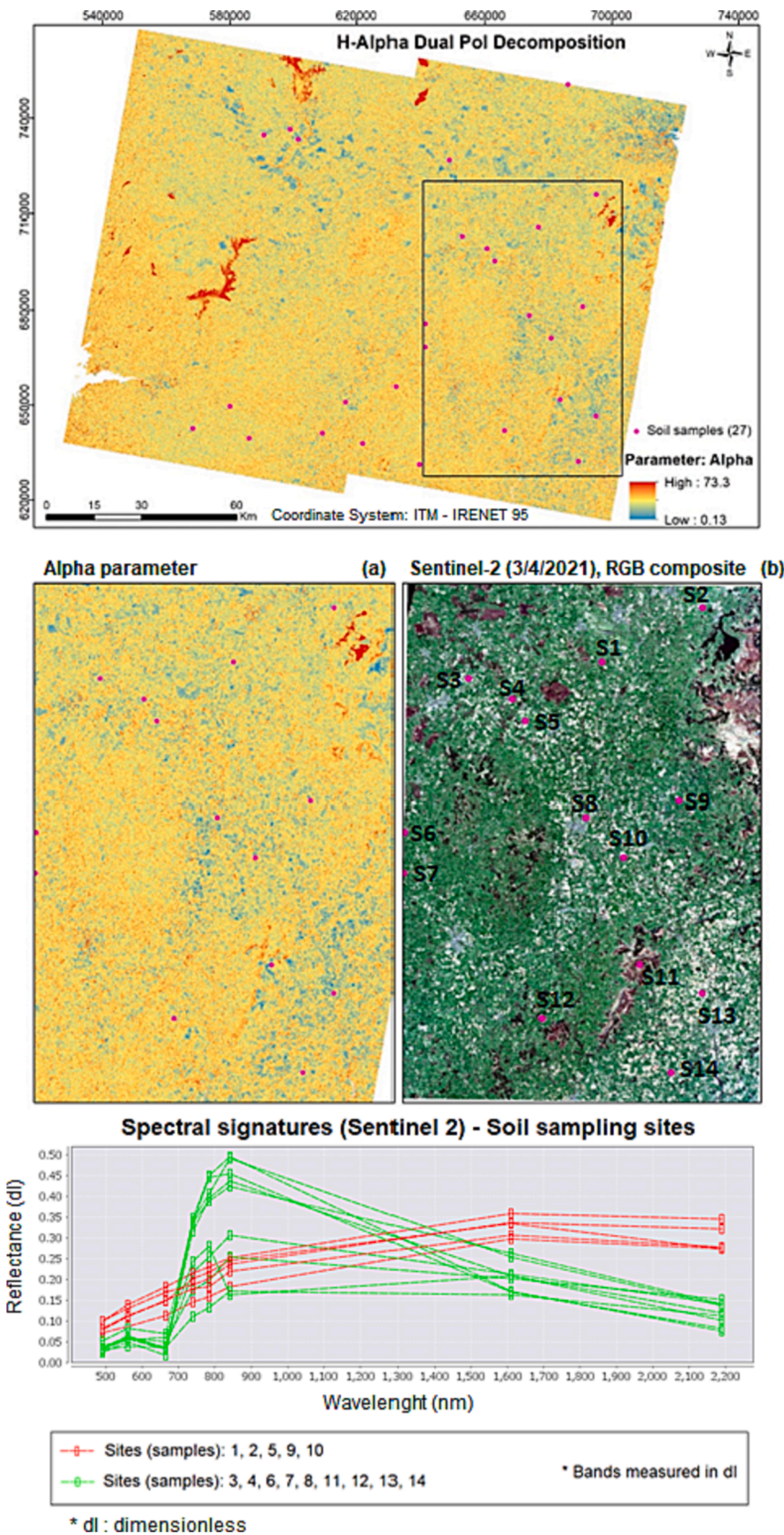
Fig. 5. A subset of Alpha of the entire scene and a comparison to the RGB composite (R4G3B2) Sentinel 2 (optical imagery) showing a correspondence between the two images (a, b). The square box in the entire scene highlights the location of the subset. Blue colour in figure "a" (surface scattering) depicts bare soil and arable lands (with very low- or more-sparse vegetation), while yellow colours (volume scattering) correspond to vegetated areas as depicted in figure "b". A spectral plot (bottom) was obtained from Sentinel 2 imagery for the selected soil sampling sites (14 points), which depicts spectral signatures of vegetation (green lines) and bare soil (red lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.13 to 73.3 (Fig. 4). For the selected soil samples (n = 27), corresponding to the surface scattering patches, the obtained alpha values ranged from 5.43 to 18.63, with a mean equal to 13.26 (Table 1). Similarly, the entropy values, indicating surface scattering, produced values ranging between 0.19 and 0.66, with a mean equal to 0.53. Alpha and Entropy followed a negative, left-tailed distribution, whereas Anisotropy exhibited a distinct positive, right-tailed distribution.

The corresponding H-Alpha plot for the study area (Fig. 4a) shows the contribution of soils and vegetation on radar backscattering. From Tables 1 and 2, the sample locations (n = 27) were consistent with the low entropy surface scattering (Alpha <= 40 and H <= 0.6). From the H-Alpha plot, the selected points (n = 27) typically lay within Z3 and Z6 (Fig. 4c), defined as brag and random surface scattering regions, respectively. Points located in Z3 and those located in Z6 in transition to Z3 indicate sparse or very low vegetation cover and soil.

The alpha values from a subset of the entire study area were visually compared to the Sentinel 2 (optical imagery) RGB composite (R4G3B2) (Fig. 5). Broadly, blue in the alpha parameter refers to surface scattering; yellow is commonly related to volume scattering; and red colour indicates a double-bounce scattering. This is also consistent with the spectral signatures obtained from Sentinel 2 (Fig. 5, bottom) of the soil sampling sites depicted in the S2-RGB composite. The spectral curves shows the signature of soil (red) and vegetation (green) for sampling sites under bare soil and vegetated soil, respectively.

A more detailed subset of alpha and Sentinel 2-RGB composite (Figs. 6 and 7) illustrates the response under different conditions of vegetation cover, on which surface scattering is depicted in areas with bare soil and very-low vegetation indicating that the decomposition method retrieved soil backscattering under that degree of vegetation cover. These figures were useful to demonstrate the method in a surrounding area of the selected points shown in Fig. 5. They have a visual purpose only, and examples from these were not included in the models.

Bare soil and low vegetation (i.e., grass) were associated with surface scattering when some pixels samples were extracted from the Alpha image (S1) and the corresponding RGB composite from S2 (Fig. 7), which could be confirmed by the H-α classification regions in the H-α plane plot. This is also consistent with the spectral signatures (Fig. 7) of a representative pixels within each polygon traced over the alpha and RGB images. The red and yellow spectral curves depict bare soil in fields (polygons) 3 and 5, respectively; green and pink curves indicate some degree of vegetation in fields 2 and 4 where soil occurs in the gaps between plants; and grey and blue spectral curves indicating the influence of dried crop residues (e.g. straw) from previous harvest in fields 1 and 6. It also worth noting that moisture content and surface roughness affect the soil spectrum, particularly in the visible spectrum (<0.750 nm).

4.1. Statistical models

The compositional LRM (Y-LRM) fitted with surface scattering as a covariate yielded an overall R2 value of 83 %. In the model without the surface scattering variable, the overall R2 value was 78.6 %. The NSE values (1:1 line) between observed and predicted samples indicated a improvement for silt and clay (Table 3). Fraction clay estimates showed the lowest agreement with the measured data in both modelling approaches (Fig. 8a and 8b). Regarding soil texture classes, the

Table 1
Descriptive statistics for the dual-pol decomposition parameters retrieved for the soil samples.

Parameter	Min.	Max.	Mean	Standard Dev.
Alpha Surface Scattering	5.43	18.63	13.26	3.16
Entropy Surface Scattering	0.19	0.66	0.53	0.11
Anisotropy Surface Scattering	0.66	0.94	0.75	0.07

Note: N = 27.

corresponding classes from the soil particle-size fractions predicted were Sand (S), Sandy Loam (SL), Loam (L), Clay Loam (CL), and Silt Loam (SIL) – Fig. 9b. The model fitted without surface scattering, resulted in more clayey classes, such as Clay (C) and Silty Clay Loam (SICL). Sand class (S) was not observed with this model.

In the tree-based models fitted with CLR transformation applied to the response variables and validated with LOOCV, the inclusion of surface scattering as a covariate improved the estimation of soil particle-size fractions. For example, the error reduced for all PSF in all models and the agreement increased, except for clay, in the RF, BART and SBART models (Table 3). The clay fraction estimates showed the lowest agreement with the measured data in all models. The sand fraction was highlighted as the optimum response variable and the LLF and SBART models exhibited the best performance (Fig. 8c). All the tree-based models yielded the same soil textural classes (Fig. 9c-9f).

For the tree-based models without log-ration transformation (non-CoDA approach), the inclusion of surface scattering as a covariate also improved the estimation of soil particle-size fractions (Table 3). In general, the models demonstrated better performance, from the NSE, RMSE and MAE metrics, between the observed and test samples, except for clay.

In an assessment of Sentinel-1, geophysical and topographical covariates for estimating topsoil particle-size fractions over the same study area, Deodoro et al. (2023) did not find a significant correlation between the radar backscatter coefficients (σ_{VV}^0, σ_{VH}^0) and the topsoil particle-size fractions for either the non-transformed or transformed soil PSF. Han et al. (2017) also found a weak correlation between radar backscatter (UAVSAR) and in situ soil PSF along the lower Mississippi river (riverside and landside). Such results indicate that soil PSF cannot be estimated directly from SAR data solely employing VV and VH backscatter values, that is, without covariates. We observed higher values for the association between SAR backscattering and SOC, however, correlations were weak with values ranging from −0.25 to 0.33 and not statistically significant (Supplementary Table 4). A possible explanation for these results, is that soil samples usually go through treatment or sample pre-treatment analytical procedures in laboratory to remove cementing materials such as organic matter or "humus", calcium carbonate and iron oxides, to obtain the particle-size fractions (Ribeiro et al., 2020).

5. Discussion

As far as we are aware, no study to date has employed dual-pol decomposition for estimating soil PSF and predicting soil texture. Hence, some comparative analyses with results from other studies employing such a technique for biophysical properties retrieving were used to underpin this discussion section.

The H-alpha Dual-Pol Decomposition method provided, to a certain extent, surface scattering over low vegetation, which may indicate underlying soil. This was observed in our study by comparing the Alpha image and the RGB composite from Sentinel 2 image. Despite using a C-band radar, which has a lower ability than an L-band radar to penetrate (low)vegetation and reach the soil (Babaeian et al., 2019). Typically, cross-polarisation (e.g. VH) is associated with multiple scattering and is useful for volume (multiple) scattering occurrence (Husman et al., 2021). Thus, it may be useful to separate soil and vegetation (Sowter, n. d.). In a study on crop growth analysis in India using S1 and target decomposition, Salma et al. (2022) used plot of classified H-α dual-pol decomposition from S1 to analyse the behaviour of the crop field with its growth, including soil plough. They found a match between the scattering shown on the H-α plane and field's temporal backscattering during growth period, which means that dual-pol decomposition, to a certain extent, dealt with gap space between soil and plants on the crop fields.

On the other hand, Ji and Wu (2015) observed that HH-HV and HV-VV SARs cannot effectively separate surface, dipole, and multiple

Table 2
The typical scattering mechanisms in an H-α feature space.

Zone	Colour	Scattering mechanism	Boundaries
3		Low entropy surface scattering	Alpha <= 40 and H <= 0.6
2		Low entropy dipole scattering	40 < Alpha <= 46 and H <= 0.6
1		Low entropy multiple scattering	Alpha > 46 and H <= 0.6
6		Medium entropy surface scattering	Alpha <= 34 and 0.6 < H <= 0.95
5		Medium entropy vegetation scattering	34 < Alpha < 46 and 0.6 < H <= 0.95
4		Medium entropy multiple scattering	Alpha >= 46 and 0.6 < H <= 0.95
8		High entropy vegetation scattering	34 < Alpha < 46 and H > 0.95
7		High entropy multiple scattering	Alpha > 46 and H > 0.95
9		High entropy surface scattering (non-feasible)	Alpha <= 34 and H > 0.95

Source (Heras, 2015). Adapted by the authors - the zone numbers were adapted to match the H-α plane plot for SAR Sentinel 1.

scattering mechanisms due to the lack of co-polarization (HH-VV), indicating that co-polarization is vital for extracting scattering mechanisms (Ji and Wu, 2015). They argue that the scatters of most zones strongly diffusing and transferring in the H-α plots for HH-HV and HV-VV SARs. Notwithstanding caveats raised by Mascolo et al. (2022) and Ji and Wu (2015), dual-pol decomposition has been employed, for example, for feature identification for classification and for biophysical parameters estimates. In order to improve results and to avoid misinterpretation on scattering and target separation, adaptation of H-alpha planes and cosine distribution for dual-pol SAR could be useful.

The pixel samples of bare soil and low vegetation (i.e. grass) that were extracted from the Alpha parameter image (from Sentinel 1) and the corresponding RGB composite from Sentinel 2 could be confirmed as being surface scattering by the H-α classification regions in the H-α plane plot. These two types of land cover are observed in the plots corresponding to the Bragg and random scattering, which are known as surface scattering regions. This can be also observed in Fig. 7.

From the results of Alpha and Entropy shown in Table 1, we can note that the samples selected (n = 27) are more related to the low entropy surface scattering (Alpha <= 40 and H <= 0.6). This can be explained by the Bragg scattering associated with surface roughness conditions, which is common in agricultural lands due to plough, hay, or straw. On the other hand, the effect of second mechanism over quasi-bare soil (i.e. sparse vegetation cover), which might be due to the Bragg scattering (also a surface scattering) or double-bounce, needs to be more investigate for dual-pol decomposition.

The dual H-α polarimetric decomposition method was useful to improve the models' performance when the samples were selected over surface scattering mechanisms through alpha information. The compositional linear regression model fitted with ILR transformation yielded better results than the model fitted without the surface scattering mechanism. The surface scattering predictor improved estimations for silt and clay on the test data. The highest magnitude NSE values for estimations in the test dataset for both scenarios were found for sand, in spite of remaining the same (NSE = 0.86). Such agreement results reflected the estimate errors, wherein RMSE and MAE values were reduced for silt and clay. Harfenmeister et al. (2021) in a study of crop growth monitoring with Sentinel 1, in Northeast Germany, identified alpha, entropy and anisotropy as important inputs in their model. They found R^2 values from single regression models for the plant height of wheat related to entropy and anisotropy with of 0.64 and 0.61, respectively. With multiple regression models of VH, VV, entropy, and alpha, R^2 values were 0.76, 0.7, 0.7 and 0.69 for plant height, wet biomass, dry biomass, and vegetation water content, respectively. These authors thereby concluded that dual-polarimetric decomposition parameters,

derived from Sentinel 1 is capable to provide meaningful input parameters for multiple regression models to improve the prediction of biophysical parameters.

Suman (2022) also employed dual pol-decomposition with ALOS-PALSAR for land feature detection and biomass estimation but over tropical deciduous forest in India. Regarding biomass results, the author obtained a correlation of 0.56 with RMSE = 29.13 t/ha and data agreement of 0.62 using exponential regression model. He concluded that decomposition parameters were more meaningful to forest structure and feature identification rather than biomass information.

Regarding the tree-based models (CLR and non-CoDA), in the modelling scenario with the surface scattering, this predictor improved the particle size-fraction estimates, particularly for sand and silt fractions regarding the NSE. The estimates error were reduced in all models for all PSF, according to RMSE and MAE values. In general, LLF and SBART models showed the best performance, while sand was the soil fraction with better estimates. Random Forest algorithm yielded the lowest evaluation metrics values amongst the tree-based models. The tree-based models with CLR transformation showed improved metrics in comparison to this modelling approach without log-ratio transformation applied (non-CoDA approach). Likewise for the compositional LRM model, a distinctive behaviour of sand and clay results was also observed with the tree-based models. Moreover, the tree-based models appeared to capture some non-linearity, particularly for clay estimation. The different performances for estimating sand, silt, and clay content (%) with the tree-based models, the could be also explained by the sample size as machine learning methods are data-based methods.

The textural classes and the pattern of the samples displayed in the soil ternary (USDA) in the test data with the Y-LRM were most consistent with the soil ternary in the measured data than those derived from the tree-based models. It is important to note that such classes are derived from the estimations of sand, silt, and clay in the statistical models, thus, they were indirectly obtained. In the soil ternary with the true (observed) data, textural classes are obtained directly from the measured data.

Despite the models' performance, the results obtained from the tree-based models fitted with a non-CoDA approach suggest that they need to be analysed with caution. An important characteristic when dealing with soil texture is the sum of the parts (components or sand, silt and clay) being equal to 100 in each sample and such a constraint has to be guaranteed when such models are employed without transformation applied to the response variables (sand, silt and clay). Moreover, in this statistical modelling approach (standard regression method), compositional data is considered as univariate data or variables, that is, they are independent of each other. This means that the predictions were
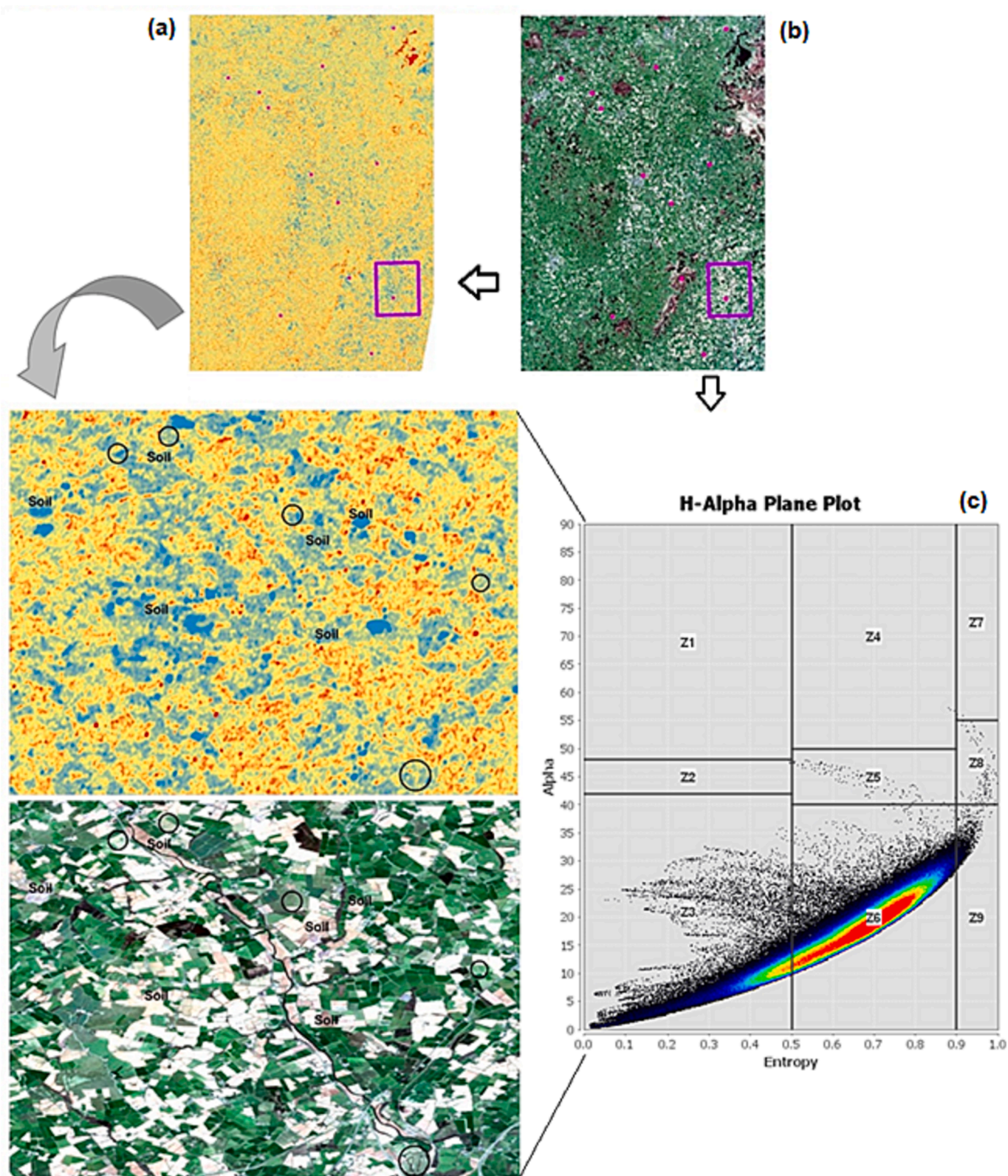
Fig. 6. A detailed subset of Alpha (a) and Sentinel 2-RGB composite (b) with the corresponding H-alpha plane plot for the subset scene (c). The figure illustrates soils under different conditions of vegetation cover. Circles on the figures were only drawn to highlight low vegetation amongst soil.

obtained separately. Filzmoser et al. (2018) highlight that the use of classical statistical methods to deal with compositional data can lead to biased results. In our work, the constraint of the total sum equal to 100 in the estimations for each sample and the response variables being treated as multivariate were taken into account.

The distinctive behaviour of soil particle-size estimations from our results suggests that sand and silt may be more related to surface scattering – sand was invariant only in the linear model. Clay did not respond to the surface scattering as sand and silt did, thereby it seems to be more related to volumetric scattering - recalling that surface and
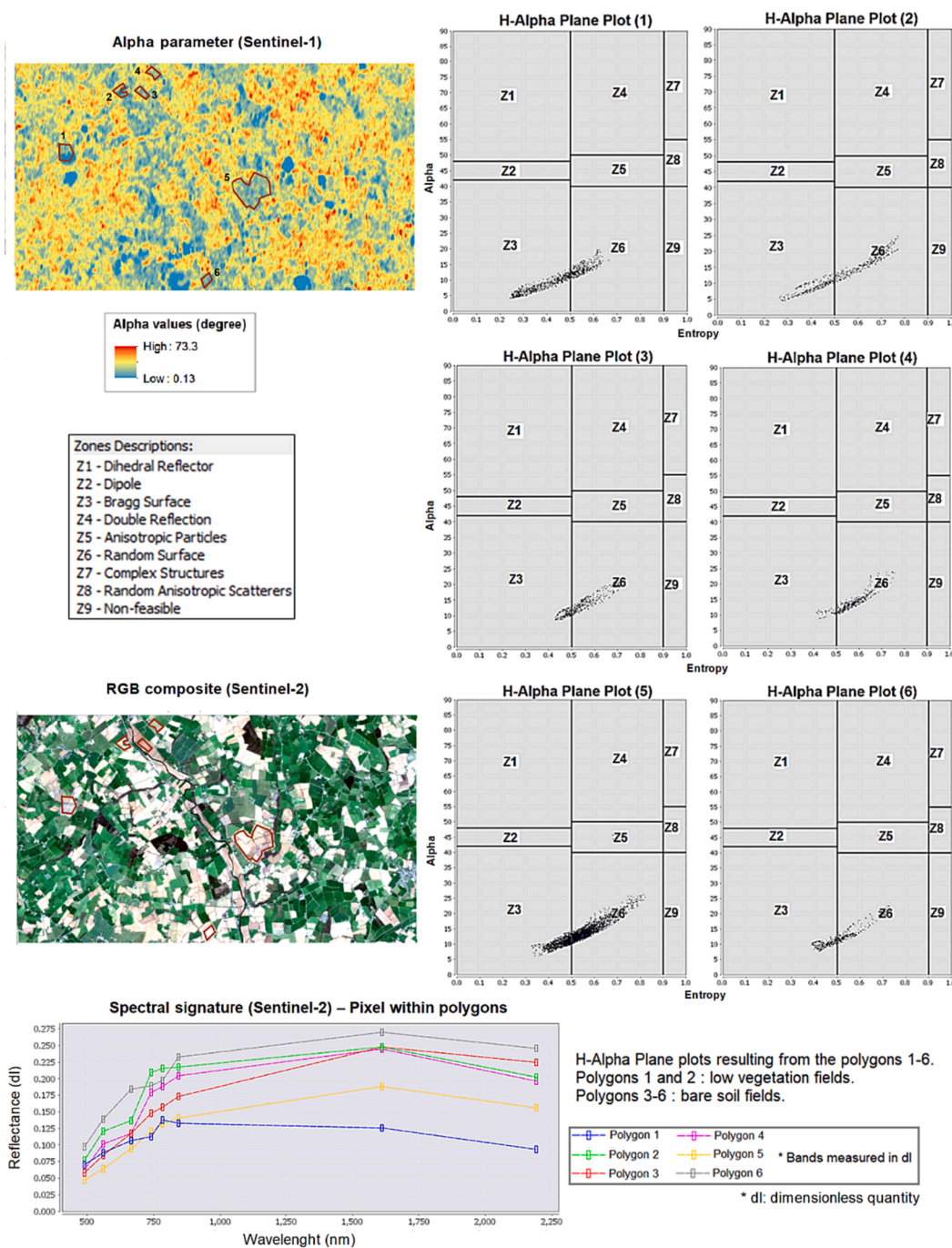
Fig. 7. H-α plane plots (right) of samples selected over bare soil and low vegetation (left) in the study area. Points (samples) located in Z3 and Z6 in the plots indicate surface scattering mechanisms as brag surface and random surface types, respectively, with low and medium entropy, respectively. In agricultural fields, brag surface may be associated with fallows while random surface could be associated with ploughed fields and fields with short grass (Anup Das, n.d: pp 27,70,77). The spectral plot on the bottom shows the spectral signature of a representative pixel within each polygon.

volume are partitions of alpha. In general, clay soil has a little more amount of plants nutrients than sandy soils and plays a vital role in holding plant nutrients and water. Notwithstanding the dry period of Sentinel 1 scenes, the returned radar beam (backscatter intensity) may have captured some remaining moisture from soils and the vegetation structure (i.e. leaves, stems, etc.), since both clay and radar signals are sensitive to moisture content. From this, the soil particle-size fractions estimations reflected in the soil textural ternary wherein more-sandy loam classes resulted from the model with the surface scattering co-variate, while more-clayey loam classes were observed from the model without that covariate.

The signal return, or backscattering coefficient (σ^0), is largely dependent on soil texture (Dobson et al., 1981) and soil moisture content because of the dielectric constant (ε) of the soil (Das and Paul, 2015). A theoretical explanation on the importance of microwave remote sensing applied to soils relies on the behaviour of soil texture in the microwave spectrum, as the distribution of particle-size fractions controls the content of free water on soils (Dobson et al., 1981). Consequently, the interpretation of SAR images relies on differences between the dielectric constant of the targets and is sensitive to the presence of water.

Regardless the statistical modelling approaches, the different patterns in the prediction of soil PSF may be related to the fact that clay

Table 3

Model evaluation metrics on test data for models validated through LOOCV method. (N = 27).

SPSF*	MODELS **WITH** SURFACE SCATTERING Y-LRM (ILR transformation)			MODELS **WITHOUT** SURFACE SCATTERING Y-LRM (ILR transformation)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	6.4	4.9	0.86	6.1	4.2	0.86
Silt	11.2	7.9	0.17	12.5	8.8	−0.33
Clay	8.0	6.0	−0.01	15.5	9.9	−1.16
SPSF*	**RF** (CLR transformation)			**RF** (CLR transformation)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	10.06	7.41	0.65	10.64	7.50	0.58
Silt	8.21	6.40	0.56	8.52	6.74	0.38
Clay	7.64	5.21	0.08	9.29	7.31	0.22
SPSF*	**BART** (CLR transformation)			**BART** (CLR transformation)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	6.83	4.83	0.84	7.51	5.27	0.79
Silt	7.29	5.81	0.65	7.95	6.35	0.46
Clay	7.03	4.87	0.22	9.11	7.09	0.25
SPSF*	**SBART** (CLR transformation)			**SBART** (CLR transformation)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	3.51	2.45	0.96	4.70	2.65	0.92
Silt	7.41	5.54	0.64	6.90	5.64	0.59
Clay	7.80	5.49	0.05	8.68	6.69	0.32
SPSF*	**LLF model** (CLR transformation)			**LLF model** (CLR transformation)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	3.84	2.97	0.95	3.70	2.57	0.95
Silt	6.61	5.56	0.71	7.10	5.53	0.57
Clay	5.86	4.16	0.42	8.36	6.33	0.37
SPSF*	**RF** (Non-CoDA approach**)			**RF** (Non-CoDA approach**)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	10.84	8.03	0.60	10.75	7.48	0.57
Silt	8.97	7.01	0.47	8.48	6.68	0.38
Clay	7.61	5.35	0.09	9.29	7.45	0.22
SPSF*	**BART** (Non-CoDA approach**)			**BART** (Non-CoDA approach**)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	6.99	5.00	0.84	8.26	5.76	0.75
Silt	7.25	5.79	0.67	8.18	6.20	0.43
Clay	7.31	5.27	0.17	9.30	7.31	0.22
SPSF*	**SBART** (Non-CoDA approach**)			**SBART** (Non-CoDA approach**)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	3.27	1.72	0.96	4.31	2.79	0.93
Silt	6.68	5.07	0.71	6.90	5.56	0.59
Clay	7.26	5.26	0.17	9.24	7.12	0.23
SPSF*	**LLF model** (Non-CoDA approach**)			**LLF model** (Non-CoDA approach**)		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Sand	4.06	3.16	0.94	4.26	3.65	0.93
Silt	6.51	5.47	0.72	7.59	5.92	0.5
Clay	5.97	4.43	0.44	8.01	6.44	0.42

RF: Random Forest; **GRF**: Generalized Random Forests; **LLF**: Local Linear Forest; **BART**: Bayesian Additive Regression Tree; **SBART**: SoftBART. Hyperparameters for each tree-based method: default. *Soil Particle-Size Fraction. ** Without log-ratio transformation.

fractions reflect more the physical and chemical soil properties, while sand and silt fractions are more related to weathering-resistant primary minerals of parent rocks (Niang et al., 2014). Moreover, in microwave remote sensing, soil texture affects the radar-backscattering coefficient since soil texture in the microwave spectrum is mirror the interplay between free water and bound water (Jackson and Schmugge, 1989; Dobson and Ulaby, 1981). Sandy soils are richer in free water than clay soils that preserve more bound water (Srivastava et al., 2006; Das and Paul, 2015). Recalling again that the Sentinel-1 dataset was acquired during a dry period (03/04/2021) inclusive of days before the Sentinel-1 passing over the study area. This is important to be considered in terms of the interaction of free water with the incident signal, which in its turn, contributes to the SAR backscatter. (Das and Paul, 2015).

5.1. Strengths and limitations of the work

Regarding the challenging task of estimating soil properties from SAR data and techniques over vegetation cover conditions, since the increased attenuation of the microwave signal over vegetated surfaces (Babaeian et al., 2019), the method applied in this work seemly ameliorates such an issue, especially when using legacy topsoil datasets as a variable response, which come from different soil surveys and measurement programmes.

This research showed that the dual H-α polarimetric decomposition method, which is typically used for classification and segmentation purposes, can also be employed for soil properties prediction from dual-pol C-band SAR such as Sentinel 1 that is freely available for users, broadening the application of this technique in microwave remote sensing studies.

Both modelling approaches showed the same pattern of results for the estimation in the test data – an improvement of silt and clay content (%) with surface scattering as one of the predictors, and sand exhibiting the highest values (although invariant with alpha-surface scattering), while clay the lowest. These findings can lead to learning more about the scattering mechanisms, soil PSF and soil texture in further works.

The performance of the models evaluated here can potentially be improved by involving other field measures such as soil moisture and the dielectric constant of soil. Moreover, a fully-pol H-alpha decomposition can be applied to an L-band SAR or even a dual-pol decomposition with a HH-VV SAR in order to compare the efficiency of the dual-pol decomposition method applied to a C-band SAR for the study area for soil PSF estimates. Could they lead to better results?

The methodological approach used in our work is not without limitations, which can be summarized in two main points. Firstly, the lack of the dielectric constant of the soil in the modelling approaches, which could be useful to analyse and exploit the relationship between sand-silt–clay and microwave backscattering coefficients (σ_{VH}^0 and σ_{VH}^0), delving more deeper into the physical scattering properties and physical soil properties. Such a data could be also useful to perform a simulation of extended-Bragg scattering scenario relating alpha, entropy and dielectric values using cosine distributions. This was carried out using full-pol data (Ponnurangam and Rao, 2017), but an adaptation of it could be useful for dual-pol VV-VH case to observe scattering angle alpha and Entropy at a particular pattern.

Secondly, the relatively small size of data obtained from the method here employed to train and test the model that might be led to an overestimation of sand % and an underestimation of clay %. Nevertheless, the results may be mirrored the true proportion among the components in the samples (compositional sand, silt, and clay fractions) rather than the sampling size.

Likewise in Ji and Wu (2015) and Mascolo et al (2022), we are aware of the inherent limitation of dual polarisation-based SAR for target decomposition, nonetheless, the results are encouraging as the models here evaluated with the surface scattering covariate produced better estimates of soil particle-size fractions than the model without surface
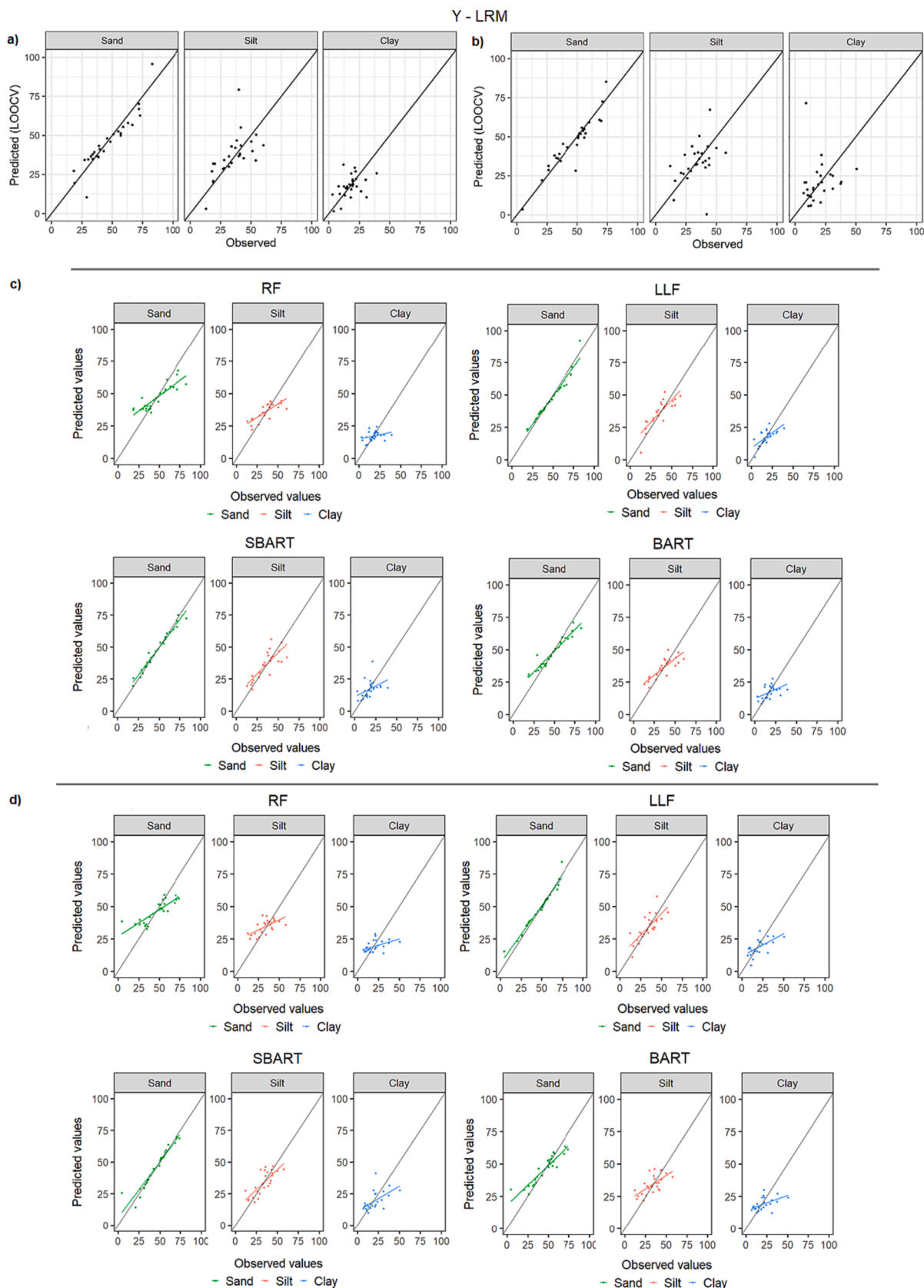
Fig. 8. Scatterplots depicting the observed and predicted data resulted from the compositional linear model and different tree-based models with the H-alpha dual-pol decomposition applied (surface scattering partition) (a, c) and without surface scattering in the model (b, d). **RF:** Random Forest; **GRF:** Generalized Random Forests; **LLF:** Local Linear Forest; **BART:** Bayesian Additive Regression Tree; **SBART:** SoftBART.
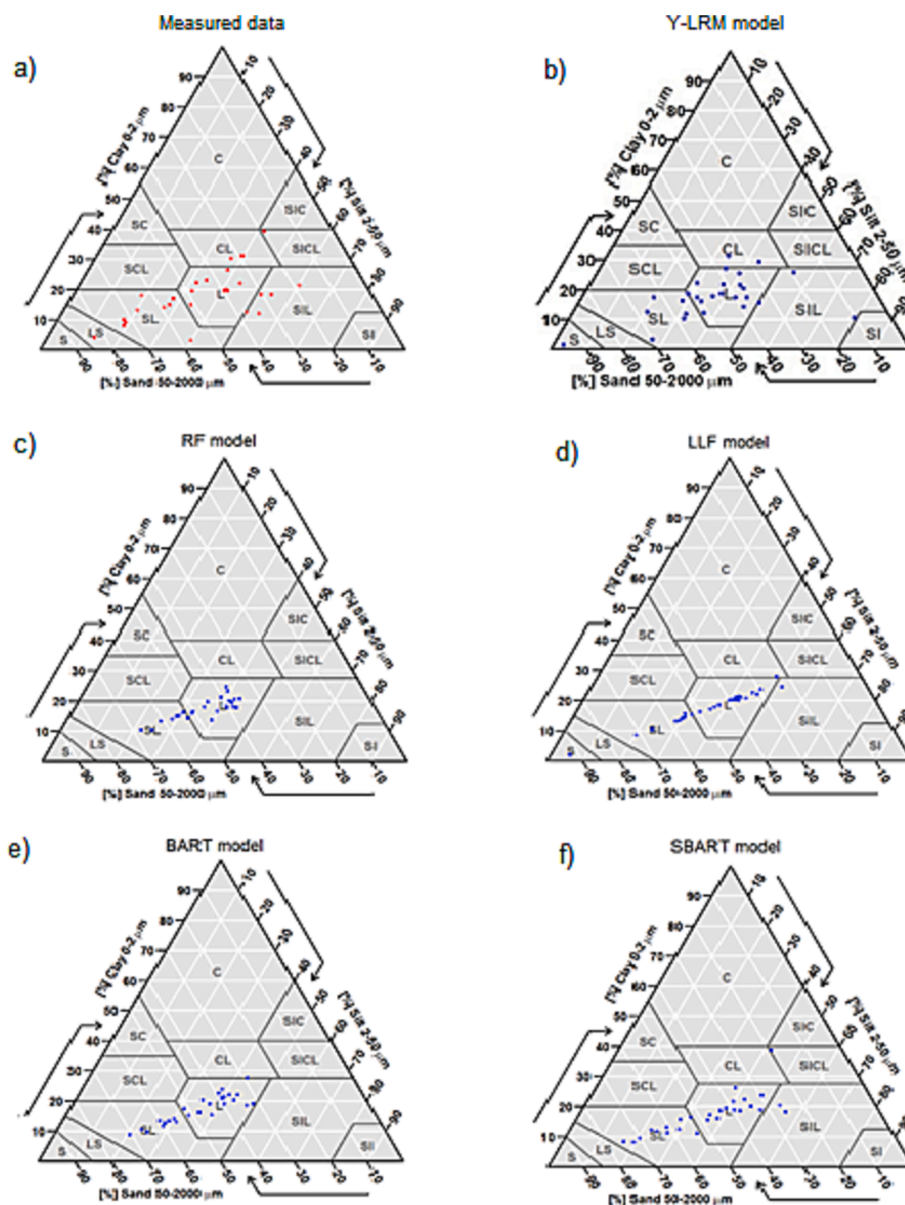
Fig. 9. Soil ternary (USDA) with the textural classes obtained from the statistical models with surface scattering data included in the models and log-ratio transformation applied to the response variables (CoDA approach). Soil ternary in figure a) refers to the classes obtained directly from the measured data. Soil textural classes inferred from the regression models are depicted in figure (b) with the compositional LRM and in figures (c-f) with the tree-model approaches.

scattering data. This indicates the potentiality of the method (H-alpha dual pol-decomposition) to obtain soil information under some degree of vegetation cover (i.e. very low-and sparse vegetation) even with C-band radar.

6. Conclusion

In this work, we tested the surface scattering mechanism, derived by H-alpha dual pol-decomposition from Sentinel 1, and additional covariates to predict percentages of compositional sand, silt, and clay (topsoil). Our results reflected the dual-pol S1 limitation, highlighted in the literature, to derive alpha values for surface scattering retrieval under vegetated soil, as the results showed a slight improvement in soil PSF estimates for the model with surface scattering values employed. Notwithstanding that, the method improved results in all the models employed in this study.

The method here applied is an empirical modelling-based approach wherein regression fits between in situ measurements – in this case study

the percentages of sand, silt, and clay – and σ°. The advantages of it are based on the fact that they are simple and straightforward, and there is no physical basis behind the model (Barrett et al., 2009). One shortcoming is that empirical models are usually only valid for the area under investigation (Barrett et al., 2009). As a result, in an example of soil texture mapping purposes, the method could be used in an area with similar characteristics that those of our study area, that is, low- and sparse vegetation (e.g. short grass, crops), bare soil, and arable- or fallow lands. Nonetheless, in light of the potential wide applicability of polarimetric decomposition methods to other than classification, the method could be employed to evaluate its usage in a more-dense vegetated area using SAR sensors built on longer wavelengths/shorter frequencies such as L-band and P-band. This could raise an issue in terms of data availability – which is not the case of Sentinel 1 since it is freely available for users. However, there is an increasing effort of space agencies to provide data freely available to everyone. For example, the Argentine SAOCOM (Satélite Argentino de Observación COn Microondas), a fully polarimetric satellite constellation of two spacecrafts that operate a

spaceborne L-band SAR system.

In the context of an agricultural field setting, one advantage of empirical modelling is the possibility to differentiate a pixel from various scattering mechanisms (Barrett et al., 2009). Our work demonstrated that, under the presence of some degree of vegetation (i.e., a low- or sparse vegetation), target decomposition theorems and Pol-SAR decomposition methods together with radar vegetation indexes such as DpRVI can be used to compensate for this. Particularly for dual-pol SAR without an HH-VV polarisation (Ji and Wu, 2015), as there is a scattering mixing in the H-alpha planes hampering an optimum separation of scatterers. This was observed in our study in the plots from Fig. 7. It is worth noting that alpha refers to a dominant scattering mechanism in a pixel. Thus, despite of being a target decomposition, the pol-decomposition technique does not separate individual targets in a pixel, but rather dominant, since the radar image pixels contain many scatterers. On the other hand, notwithstanding constraints about dual-pol decomposition, the H-alpha Dual Pol-Decomposition method with S1-C band SAR was able to capture the gaps between soil and vegetation through the dominant targets within a pixel (10 m spatial resolution). This was also demonstrated in Figs. 6 and 7, which was also corroborated by spectral signatures of a representative pixel within each polygon. From this, dual-pol decomposition remains a potential method for retrieving soil properties and biophysical parameters when employed with caution and meticulousness.

Digital soil mapping generally relies on empirical models to integrate ancillary data influenced by topographical variables derived from a digital elevation model to soil particle-size fractions to understand distribution patterns of soil PSF across an area. (Mondejar and Tongco, 2019). In this context, methods outlined here have the advantage of accounting for the requirements of compositional data such as soil texture. Moreover, regarding the laborious soil sampling associated with soil surveys, this work highlights the possibility of combing model based approaches with more limited in situ sampling methods – especially for sand and silt – for a sample size of 27 observations even splitting data into training and testing.

Thus, the exploratory assessment of the methods carried out in this work is a starting point for the development of other case studies to ameliorate predictions of soil particle-size fractions and soil texture over low-and short vegetated areas, going beyond the traditional polarimetric decomposition technique applied for classification and segmentation of landscapes or for vegetation studies.

Future works on this application – soil PSF and soil texture predictions with SAR techniques – may be beneficial by employing different backscattering coefficients – particularly gamma nought –, other speckle filtering methods, cosine distributions, and soil-specific covariates related to SAR such as dielectric constant, soil roughness, and soil moisture. Moreover, in light of the endeavour of the Sentinel satellite constellation and its benefits (e.g., free availability, temporal-radiometric-spatial resolutions, and orbit stability), joining Sentinel 1 and Sentinel 2 products and their techniques can be also beneficial. For example, combining H-alpha dual-pol decomposition with spectral unmixing method (i.e. spectral mixture analysis using endmember values) to extract values to serve as input in a statistical model, as they are pixel-based analyses. A pure pixel or endmember is considered as such when a pixel contains a reflectance of one land cover/use feature. The aim of the spectral unmixing method is roughly the same as that of target polarimetric decomposition, that is, to distinguish the contributions of the targets in a single pixel since it is not pure due to the randomness of scatterings.

Funding

CRediT authorship contribution statement

Sandra Cristina Deodoro: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Rafael de Andrade Moral:** Formal analysis, Methodology, Writing – review & editing. **Reamonn Fealy:** Investigation, Resources, Writing – review & editing. **Tim McCarthy:** Resources, Writing – review & editing. **Rowan Fealy:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2024.103742.

References

Abdel-Hamid, A., Dubovyk, O., El-Magd, I., Menz, G., 2018. Mapping Mangroves extents on the red sea coastline in Egypt using polarimetric SAR and high-resolution optical remote sensing data. Sustainability. https://doi.org/10.3390/su10030646.

Aitchison, J., 1982. The statistical analysis of compositional data (with discussion). J. R. Stat. Soc. Ser. B Stat. Methodol 44 (2), 139–177.

Aitchison, J. (2005). A Concise Guide to Compositional Data Analysis. Available at: https://eprints.gla.ac.uk/259608/ (accessed 19 Dez 2022).

Amirian-Chakan, A., Minasny, B., Taghizadeh-Mehrjardi, R., Akbarifazli, R., Darvishpasand, Z., Khordehbin, S., 2019. Some practical aspects of predicting texture data in digital soil mapping. Soil Tillage Res. 194, e104289.

Anup Das (n.d.). Polarimetric_SAR_Data_Analysis_and_Applications. Available at https://vedas.sac.gov.in/vedas/downloads/ertd/SAR/L_5_Polarimetric_SAR_Data_Analysis_and_Applications_Dr_Anup_Das.pdf (accessed 25 July 2023).

Azizi, K., Garosi, Y., Ayoubi, S., Tajik, S., 2023. Integration of Sentinel-1/2 and topographic attributes to predict the spatial distribution of soil texture fractions in some agricultural soils of western Iran. Soil Tillage Res. 229, 105681 https://doi.org/10.1016/j.still.2023.105681.

Babaeian, E., Sadeghi, M., Jones, S.B., Montzka, C., Vereecken, H., Tuller, M., 2019. Ground, proximal, and satellite remote sensing of soil moisture. Rev. Geophys. 57 https://doi.org/10.1029/2018RG000618.

Baghdadi, N., Choker, M., Zribi, M., Hajj, M.E., Paloscia, S., Verhoest, N.E., Lievens, H., Baup, F., Mattia, F., 2016. A new empirical model for radar scattering from bare soil surfaces. Remote Sens. 8 (11), 920. https://doi.org/10.3390/rs8110920.

Ballabio, C., Panagos, P., Montanarella, L., 2016. Mapping topsoil physical properties at European scale using the LUCAS database (2016). Geoderma 261, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006.

Barrett, B.W., Dwyer, E., Whelan, P., 2009. Soil moisture retrieval from active spaceborne microwave observations: an evaluation of current techniques. Remote Sens. 1, 210–242. https://doi.org/10.3390/rs1030210.

Barrett, B., Dwyer, N., Whelan, P., & Bartlett, D. (2007). Soil moisture determination in southern Ireland using an ASAR time series. In: Proceedings of Envisat Symposium, April 2007. Montreux, Switzerland: European Space Agency (Special Publication)

ESA SP-636. Available at: https://earth.esa.int/envisatsymposium/proceedings/sessions/4D1/460228BB.pdf (accessed 06 Jun 2021).

Bhogapurapu, N., Dey, S., Mandal, D., Bhattacharya, A., Karthikeyan, L., McNairn, H., Rao, Y., 2022. Soil moisture retrieval over croplands using dual-pol L-band GRD SAR data. Remote Sens. Environ. 271, 112900.

Bousbih, S., Zribi, M., Pelletier, C., Gorrab, A., Lili-Chabaani, Z., Baghdadi, N., Aissa, N. B., Mougenot, B., 2019. Soil texture estimation using radar and optical data from sentinel-1 and sentinel-2. Remote Sens. 11 (13), 1520. https://doi.org/10.3390/rs11131520.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Chappell, M.A., Seiter, J.M., West, H.M., Durham, B.D., Porter, B.E., Price, C.L., 2019. Building geochemically based quantitative analogies from soil classification systems using different compositional datasets. PLoS One 14 (2), e0212214.

Chipman, H.A., George, E.I., McCulloch, R.E., 2010. Bart: Bayesian additive regression trees. Ann. Appl. Stat. 4 (1), 266–298. https://www.jstor.org/stable/27801587.

Cisty, M., Soldanova, V., Cyprich, F. (2019). Random forest based reclassification of soil texture for hydrological modelling. Geophysical Research Abstracts, 21, EGU2019-14320. Available at https://meetingorganizer.copernicus.org/EGU2019/EGU2019-14320.pdf (accessed 22 June 2023).

Cloude, S.R., Pottier, E., 1996. A review of target decomposition theorems in radar polarimetry. IEEE Trans Geosci Remote 34 (2), 498–518. https://doi.org/10.1109/36.485127.

Cloude, S.R., Pottier, E., 1997. An entropy based classification scheme for land applications of polarimetric SAR. IEEE Trans. Geosci. Rem. 35 (1), 68–78. https://doi.org/10.1109/36.551935.

Coblinski, J.A., Inda, A.V., Demattê, J.A.M., Dotto, A.C., Gholizadeh, A., Giasson, E., 2021. Identification of minerals in subtropical soils with different textural classes by VIS-NIR-SWIR reflectance spectroscopy. Catena (amst) 203, e105334.

Comas-Cufí, M., Thió-Henestrosa, S., 2011. CoDaPack v2.03 (Version 2.03). IMAE-UDG. http://imae.udg.edu/codapack/.

Creamer, R. E., Simo, I., O'Sullivan, L., Reidy, B., Schulte, R. P. O., Fealy, R. M. (2016). Irish Soil Information System: Soil Property Maps. Report N. 204. EPA Research Programme 2014–2020. Teagasc, Wexford, Ireland.

Das, K., Paul, P.K., 2015. Present status of soil moisture estimation by microwave remote sensing. Cogent Geosci. 1 (1) https://doi.org/10.1080/23312041.2015.1084669.

Dave, R.B., Saha, K., Kushwaha, A., Vithalpura, M., Nidhin, P., Murugesan, A., 2023. Analysing the potential of polarimetric decomposition parameters of Sentinel–1 dual-pol SAR data for estimation of rice crop biophysical parameters. J. Agrometeorol. 25 (1), 105–112. https://doi.org/10.54386/jam.v25i1.2039.

De Zan, F., Parizzi, A., Prats-Iraola, P., López-Dekker, P., 2014. A SAR interferometric model for soil moisture. IEEE Trans. Geosci. Remote 52 (1), 418–425. https://doi.org/10.1109/TGRS.2013.2241069.

Deodoro, S.C., Moral, R. de A.; Fealy; R.; McCarthy, T., & Fealy, R. (2023). An assessment of Sentinel 1 SAR, geophysical and topographical covariates for estimating topsoil particle size fractions. Eur. J. Soil Sci., 74(5), e13414, DOI: 10.1111/ejss.13414.

Dobson, M.C., Ulaby, F.T., 1981. Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: Part III-soil tension. IEEE Trans. Geosci. Remote Sens. 19, 51–61. https://doi.org/10.1109/TGRS.1981.350328.

Dobson, M.C., Ulaby, F.T., Hallikainen, M.T., El-rayes, M.A., 1985. Microwave dielectric behavior of wet soil, Part II: dielectric mixing models. IEEE Trans. Geosci. Remote Sens. Vol. GE-23, 35–46. https://doi.org/10.1109/TGRS.1985.289498.

Domenech, M.B., Amiotti, N.M., Costa, J.L., Castro-Franco, M., 2020. Prediction of topsoil properties at field-scale by using C-band SAR data. Int. J. Appl. Earth Obs. Geoinf. 93, 102197 https://doi.org/10.1016/j.jag.2020.102197.

dos Santos, J. R., Narvaes, I. S., S., I., A. Graca, P. M. L. A., Gonçalves, F. G. (2009). Polarimetric Responses and Scattering Mechanisms of Tropical Forests in the Brazilian Amazon. Advances in Geoscience and Remote Sensing. InTech. https://doi.org/10.5772/8340.

Dotto, A.C., Demattè, J.A.M., Rossel, R.A.V., Rizzo, R., 2020. Soil environment grouping system based on spectral, climate, and terrain data: a quantitative branch of soil series. Soil 6, 163–177. https://doi.org/10.5194/soil-6-163-2020.

Dou, Q., Xie, Q., Peng, X., Lai, K., Wang, J., Lopez-Sanchez, J.M., Shang, J., Shi, H., Fu, H., Zhu, J., 2022. Soil moisture retrieval over crop fields based on two-component polarimetric decomposition: a comparison of generalized volume scattering models. J. Hydrol. 615,A https://doi.org/10.1016/j.jhydrol.2022.128696.

Engelbrecht, J., Theron, A., Vhengani, L., Kemp, J., 2017. A simple normalized difference approach to burnt area mapping using multi-polarisation C-Band SAR. Remote Sens. 2017 (9), 764. https://doi.org/10.3390/rs9080764.

European Space Agency-ESA (n.d.). Sentinel Online. Interferometric Wide Swath. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/interferometric-wide-swath (accessed 28 October 2023).

European Space Agency-ESA (2007). ASAR Product Guide Book, Chapter 1 [on line]. https://earth.esa.int/eogateway/documents/20142/37627/ASAR-Product-Handbook.pdf (accessed 25 October 2023).

Fealy, R., Bruyère, C., Duffy, C.(2011). Regional Climate Model Simulations for Ireland for the 21st Century. EPA Research Report 2011-CCRP-MS-2.2 Environmental Protection Agency, Wexford, Ireland. Available at https://www.epa.ie/publications/research/climate-change/Research_Report_244.pdf (accessed 3 November 2023).

Filzmoser, P., Hron, K., Templ, M., 2018. Applied Compositional Data Analysis With Worked Examples in R., first ed. Springer Cham, Switzerland Doi: 10.1007/978-3-319-96422-5.

Fletcher, R., 2022. Temporal comparisons of apparent electrical conductivity: a case study on clay and loam soils in Mississippi. Agric. Sci. 13, 936–946. https://doi.org/10.4236/as.2022.138058.

Freeman, A., Durden, S.L., 1998. A three-component scattering model for polarimetric SAR data. IEEE Trans Geosci Remote 36 (3), 963–973. https://doi.org/10.1109/36.673687.

Freeman, A. (nd). Radiometric calibration of SAR image data. International Society for Photogrametry and Remote Sensing. ISPRS Proceedings, XXIX, part 1, 2012. Available at https://www.isprs.org/proceedings/xxix/congress/part1/212_xxix-part1.pdf (accessed 23 October 2023).

Friedberg, R., Tibshirani, J., Athey, S., Wager, S., 2021. Local Linear Forests. J. Comput. Graph. Stat. 30 (2), 503–517. https://doi.org/10.1080/10618600.2020.1831930.

Gholizadeha, A., Žižala, D., Aberioon, M.C., Borůvka, L., 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. Remote Sens. Environ. 218, 89–103. https://doi.org/10.1016/j.rse.2018.09.015.

Guerriero, L., Pierdicca, N., Pulvirenti, L., Ferrazzoli, P., 2013. Use of satellite radar bistatic measurements for crop. Rem. Sens. 5 (2), 864–890. https://doi.org/10.3390/rs5020864.

Gururaj, P., Umesh, P., Shetty, A., 2019. Assessment of spatial variation of soil moisture during maize growth cycle using SAR observations. In: Proc. SPIE 11149, Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI, 1114916, Doi: 10.1117/12.2532953.

Halbgewachs, M., Wegmann, M., da Ponte, E., 2022. A spectral mixture analysis and landscape metrics based framework for monitoring spatiotemporal forest cover changes: a case study in mato Grosso. Brazil. Remote Sens. 14, 1907. https://doi.org/10.3390/rs14081907.

Hallikainen, M.T., Ulaby, F.T., Dobson, M.C., El-rayes, M.A., Wu, L.K., 1985. Microwave dielectric behavior of wet soil-Part 1: Empirical models and experimental observations. IEEE Trans. Geosci. Remote Sens. 23, 25–34. https://doi.org/10.1109/TGRS.1985.289497.

Han, D., Vahedifard, F., Aanstoos, J.V., 2017. Investigating the correlation between radar backscatter and in situ soil property measurements. Int. J. Appl. Earth Obs. Geoinf. 57, 136–144. https://doi.org/10.1016/j.jag.2016.12.018.

Harfenmeister, K., Itzerott, S., Weltzien, C., Spengler, D., 2021. Agricultural monitoring using polarimetric decomposition parameters of sentinel-1 Data. Remote Sens. 13 (4), 575. https://doi.org/10.3390/rs13040575.

Heras, A. B. I. O. (2015). Decomposition and unsupervised segmentation of dual-polarized polarimetric SAR data using fuzzy entropy and coherency clustering method. Master dissertation in Telecommunication Engineering. Universitat Autònoma de Barcelona. Escola d'Enginyeria. Available at https://ddd.uab.cat/record/133501?ln=en (accessed 14 Dec 2023).

Husman, S. de R., Van der Sanden, J.J., Lhermitte, S., Eleveld, M.A., 2021. Integrating intensity and context for improved supervised river ice classification from dual-pol Sentinel-1 SAR data. Int. J. Appl. Earth Obs. Geoinf. 101 https://doi.org/10.1016/j.jag.2021.102359.

Imperatore, P., Di Martino, G., 2023. SAR radiometric calibration based on differential geometry: from theory to experimentation on SAOCOM Imagery. Remote Sens. 15, 1286. https://doi.org/10.3390/rs15051286.

Jackson, T.J., Schmugge, T.J., 1989. Passive microwave remote-sensing system for soil moisture: Some supporting research. IEEE Trans Geosci Remote 27 (2), 225–235. https://doi.org/10.1109/36.20301.

Jackson, Thomas. J., 1987. Effects of soil properties on microwave dielectric constants. Transp Res Rec, 1119, 126-131. Available at https://onlinepubs.trb.org/Onlinepubs/trr/1987/1119/1119-016.pdf (accessed 2 November 2023).

Jagdhuber, T., Hajnsek, I., Bronstert, A., Papathanassiou, K.P., 2013. Soil moisture estimation under low vegetation cover using a multiangular polarimetric decomposition. IEEE Trans. Geosci. Remote 51 (4), 2201–2215. https://doi.org/10.1109/TGRS.2012.2209433.

Jagdhuber, Thomas, 2012. Soil Parameter Retrieval under Vegetation CoverUsing SAR Polarimetry. Unpublished PhD thesis. University of Potsdam, Potsdam, Germany. Available at: https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/deliver/index/docId/5894/file/jagdhuber_diss.pdf (accessed 3 August 2023).

Ji, K., Wu, Y., 2015. Scattering mechanism extraction by a modified cloude-pottier decomposition for dual polarization SAR. Remote Sens. 7 (6), 7447–7470. https://doi.org/10.3390/rs70607447.

Katebikord, A., Sadeghi, S.H., Singh, V.P., 2022. Spatial modeling of soil organic carbon using remotely sensed indices and environmental field inventory variables. Environ. Monit. Assess. 194, 152.

Khajehzadeh, M., Afzali, S.F., Honarbakhsh, A., et al., 2022. Remote sensing and GIS-based modeling for predicting soil salinity at the watershed scale in a semi-arid region of southern Iran. Arab. J. Geosci. 15, 423. https://doi.org/10.1007/s12517-022-09762-4.

Kiely, G., Ms. McGoff, N.M., Eaton, J.M., Xu, X., Leahy, P., Carton, O., 2009. SoilC - Measurement and Modelling of Soil Carbon Stocks and Stock Changes in Irish Soils. EPA STRIVE Programme 2001-2007. STRIVE Report. SoilC Final Report – June 12, 2009. Available at: https://www.ucc.ie/en/media/research/hydromet/EPAkielyReportSoilC.2009.pdf (accessed 3 August 2023).

Koppe, W., Gnyp, M.L., Hütt, C., Yao, Y., Miao, Y., Chen, X., Bareth, G., 2013. Rice monitoring with multi-temporal and dual-polarimetric TerraSAR-X data. Int. J. Appl. Earth Obs. Geoinf. 21, 568–576. https://doi.org/10.1016/j.jag.2012.07.016.

Lee, J-S., Pottier, E., 2009. Polarimetric Radar Imaging: From Basics to Applications, first ed. CRC Press, New York. Doi: 10.1201/9781420054989.

Linero, A.R., Yang, Y., 2018. Bayesian tree ensembles that adapt to smoothness and sparsity. J. R. Stat. Soc., Series B 80 (5), 1087–1110. https://doi.org/10.1111/rssb.12293.

Loosvelt, L., Vernieuwe, H., Pauwels, V.R.N., De Baets, B., Verhoest, N.E.C., 2013. Local sensitivity analysis for compositional data with application to soil texture in

hydrologic modelling. Hydrol. Earth Syst. Sci. 17 (461–478), 2013. https://doi.org/10.5194/hess-17-461-2013.

Magagi, R., Jammali, S., Goïta, K., Wang, H., Colliander, A., 2022. Potential of L- and C-Bands Polarimetric SAR Data for Monitoring Soil Moisture over Forested Sites. Remote Sens. 14, 5317. https://doi.org/10.3390/rs14215317.

Mandal, D., et al., 2020. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. Remote Sens. Environ. 247, 111954 https://doi.org/10.1016/j.rse.2020.111954.

Manolakis, D., Lockwood, R., Cooley, T., 2016. Spectral Mixture Analysis. In: Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms. Cambridge University Press, Cambridge, pp. 443–493. https://doi.org/10.1017/CBO9781316017876.010.

Marzahn, P., Meyer, S., 2020. Utilization of multi-temporal microwave remote sensing data within a Geostatistical regionalization approach for the derivation of soil texture. Remote Sens 12 (16), 2660. https://doi.org/10.3390/rs12162660.

Mascolo, L., Cloude, S.R., Lopez-Sanchez, J.M., 2022. Model-based decomposition of dual-Pol SAR data: application to sentinel-1. IEEE Trans. Geosci. Rem. 60 (1–19), e5220119.

Met Éireann-The Irish Meteorological Service., 2021. Climate Statement for March 2021. https://www.met.ie/climate-statement-for-march-2021 (accessed 19 June 2023).

Metternicht, G.I., Zinck, J.A., 2003. Remote sensing of soil salinity: potentials and constraints. Remote Sens. Environ. 85 (1), 1–20. https://doi.org/10.1016/S0034-4257(02)00188-8.

Mirzaeitalarposhti, R., Shafizadeh-Moghadam, H., Taghizadeh-Mehrjardi, R., Demyan, M.S., 2022. Digital soil texture mapping and spatial transferability of machine learning models using sentinel-1, Sentinel-2, and terrain-derived covariates. Remote Sens. 14, 5909. https://doi.org/10.3390/rs14235909.

Mondejar, J.P., Tongco, A.F., 2019. Estimating topsoil texture fractions by digital soil mapping - a response to the long outdated soil map in the Philippines. Sustain. Environ. Res. 29, e31.

Morais, Thomas-Agnan, 2021. Impact of covariates in compositional models and simplicial derivatives. Austrian J. Stat. 50, 1–15. https://doi.org/10.17713/ajs.v50i2.1069.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual model. Part 1-A discussion of principles. J. Hydrol. 10 (3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nasirzadehdizaji, R., Sanli, F.B., Abdikan, S., Cakir, Z., Sekertekin, A., Ustuner, M., 2019. Sensitivity analysis of multi-temporal sentinel-1 SAR parameters to crop height and canopy coverage. Appl. Sci. 9 (4), 655. https://doi.org/10.3390/app9040655.

Navacchi, C., Cao, S., Bauer-Marschallinger, B., Snoeij, P., Small, D., Wagner, W., 2023. Utilising Sentinel-1's orbital stability for efficient pre-processing of radiometric terrain corrected gamma nought Backscatter. Sensors 23, 6072. https://doi.org/10.3390/s23136072.

Niang, M.A., Nolin, M.C., Jégo, G., Perron, I., 2014. Digital mapping of soil texture using RADARSAT-2 polarimetric synthetic aperture radar data. Soil Sci. Soc. Am. J. 78, 673–684. https://doi.org/10.2136/sssaj2013.07.0307.

Odeh, I.O.A., Todd, A.J., Triantafilis, J., 2003. Spatial prediction of soil particle-size fractions as compositional data. Soil Sci. 168 https://doi.org/10.1097/01.ss.0000080335.10341.23.

Pawlowsky-Glahn, V., Egozcue, J. J. & Tolosona-Delgado, R. (2015). Modeling and Analysis of Compositional Data, Wiley, 2015, 272 p.

Periasamy, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1. Remote Sens. Environ. 217, 537–549. https://doi.org/10.1016/j.rse.2018.09.003.

Petropoulos, G.P., Ireland, G., Barrett, B., 2015. Surface soil moisture retrievals from remote sensing: Current status, products & future trends. Phys. Chem. Earth 83–84, 36–56. https://doi.org/10.1016/j.pce.2015.02.009.

Ponnurangam, G.G., Rao, Y.S., 2017. Evaluation of different orientation angle distributions within the X-Bragg scattering model for bare soil moisture estimation. Int. J. Remote Sens. 38 (15), 4379–4395. https://doi.org/10.1080/01431161.2017.1320447.

Pratola, C., Barrett, B., Gruber, A., Kiely, G., Dwyer, E., 2014. Evaluation of a global soil moisture product from finer spatial resolution SAR data and ground measurements at Irish sites. Remote Sens 6, 8190–8219. https://doi.org/10.3390/rs6098190.

Qu, Y.; Zhao, W.; Yuan, Z.; Chen, J.(2020). Crop Mapping from Sentinel-1 Polarimetric Time-Series with a Deep Neural Network. Remote Sens. 2020, 12, 2493. Doi: 10.3390/rs12152493.

Qu, J., Qiu, X., Wang, W., Wang, Z., Lei, B., Ding, C.A., 2022. Comparative study on classification features between high-resolution and polarimetric SAR images through unsupervised classification methods. Remote Sens. 14, 1412. https://doi.org/10.3390/rs14061412.

Read, C.F., Duncan, D.H., Ho, C.Y.C., White, M., Vesk, P.A., 2018. Useful surrogates of soil texture for plant ecologists from airborne gamma-ray detection. Nat. Ecol. Evol. 8 (4), 1974–1983. https://doi.org/10.1002/ece3.3417.

Ribeiro E., Batjes, N. H., van Oostrum, A. J. M. (2020). World Soil Information Service (WoSIS)-Towards the standardization and harmonization of world soil profile data. Procedures manual 2020, Report 2020/01, ISRIC - World Soil Information, Wageningen. https://doi.org/10.17027/isric-wdc-2020-01.

Salma, S., Keerthana, N., Dodamani, B.M., 2022. Target decomposition using dual-polarization sentinel-1 SAR data: Study on crop growth analysis. Remote Sens. Appl.: Soc. Environ. 28, e100854.

Schmugge, T.J., 1980. Effects of texture on microwave emission from soils. IEEE Trans. Geosci. Remote Sens. Vol. GE-18, 353–361.

Small, D., 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49, 3081–3093.

Somers, B., Asner, G.P., Tits, L., Coppin, P., 2016. Endmember variability in spectral mixture analysis: a review. Remote Sens. Environ. 115 (7), 1603–1616. https://doi.org/10.1016/j.rse.2011.03.003.

Sowter, Andrew (n.d.). Introduction to Radar. Landmap Geoknowledge. Available at https://learningzone.rspsoc.org.uk/index.php/Learning-Materials/Radar-Imaging/Image-Interpretation-Polarisation (accessed 3 August 2023).

Srivastava, H.S., Patel, P., Navalgund, R.R., 2006. How far SAR has fulfilled its expectation for soil moisture retrieval? Proc. SPIE, Microwave Remote Sensing of the Atmosphere and Environment V, e641001, Doi: 10.1117/12.693946.

Suman, S., 2022. H/A/α polarimetric decomposition of dual polarized alos palsar for efficient land feature detection and biomass estimation over tropical deciduous forest. Geogr. Environ. Sustain. 3 (15), 37–46. https://DOI-10.24057/2071-9388-2021-095.

Todorov, V., 2021. Monitoring robust estimates for compositional data. Austrian J. Stat. 50, 16–37. http://www.ajs.or.at/doi:10.17713/ajs.v50i2.1067.

van Wesemael, B., Chabrillat, S., Sanz Dias, A., Berger, M., Szantoi, Z., 2023. Remote sensing for soil organic carbon mapping and monitoring. Remote Sens. 15, 3464. https://doi.org/10.3390/rs15143464.

Walsh, S. (2012). A summary of climate averages for Ireland 1981-2010. Climatological Note no.14. Met Éireann. Dublin, May 2012. Available at: https://www.met.ie/climate-ireland/SummaryClimAvgs.pdf (accessed 22 March 2023).

Wang, H., Magagi, R., Goita, K., 2017. Comparison of different polarimetric decompositions for soil moisture retrieval over vegetation covered agricultural area. Remote Sens. Environ. 199, 120–136. https://doi.org/10.1016/j.rse.2017.07.008.

Wang, J.R., Schmugge, T.J., 1980. An Empirical Model for the Complex Dielectric Permittivity of Soils as a Function of Water Content. IEEE Trans. Geosci. Remote Sens. GE-18, 288–295. https://doi.org/10.1109/TGRS.1980.350304.

Wang, Z., Shi, W., 2017. Mapping soil particle-size fractions: A comparison of compositional kriging and log-ratio kriging. J. Hydrol. 546, 526–541. https://doi.org/10.1016/j.jhydrol.2017.01.029.

Wang, Z., Shi, W., 2018. Robust variogram estimation combined with isometric log-ratio transformation for improved accuracy of soil particle-size fraction mapping. Geoderma 324, 56–66. https://doi.org/10.1016/j.geoderma.2018.03.007.

Weiss, S., Xu, Z.Z., Peddada, S., et al., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5, 27. https://doi.org/10.1186/s40168-017-0237-y.

Woźniak, E., Rybicki, M., Kofman, W., Aleksandrowicz, S., Wojtkowski, C., Lewiński, S., Bojanowski, J., Musiał, J., Milewski, T., Slesiński, P., Łączyński, A., 2022. Multi-temporal phenological indices derived from time series Sentinel-1 images to country-wide crop classification. Int. J. Appl. Earth Obs. Geoinf. 107, e102683.

Yang, C., Yang, L., Zhang, L., Zhou, C., 2023. Soil organic matter mapping using INLA-SPDE with remote sensing based soil moisture indices and Fourier transforms decomposed variables. Geoderma 437, 116571. https://doi.org/10.1016/j.geoderma.2023.116571.

Zhang, M., Shi, W. (2019). Systematic comparison of five machine-learning methods in classification and interpolation of soil particle size fractions using different transformed data. Hydrol Earth Syst Sci.. Discussions [preprint]. Doi: 10.5194/hess-2018-584, 2019.

Zhang, Z., Jiang, X., Wang, M., Liu, X., Wang, Q., Li, R., 2022. A phase-decomposition-based polarimetric coherence optimization method. Int. J. Appl. Earth Obs. Geoinf. 110, e102771.

Zhang, Y., Yang, B., Liu, X., Wang, C., 2017. Estimation of rice grain yield from dual-polarization Radarsat-2 SAR data by integrating a rice canopy scattering model and a genetic algorithm. Int. J. Appl. Earth Obs. Geoinf. 57, 75–85. https://doi.org/10.1016/j.jag.2016.12.014.

Žížala, D., Minařík, R., Zádorová, T., 2019. Soil organic carbon mapping using multispectral remote sensing data: prediction ability of data with different spatial and spectral resolutions. Remote Sens. 11, 2947. https://doi.org/10.3390/rs11242947.