Sentinel-1 time series applications over agricultural fields: proposal, evaluation and comparison of different methodologies

María Arias Cuenca

November 2022

Supervisors:

Jesús Álvarez Mozos

Miguel Ángel Campo Bescós

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Thesis submitted in fulfillment of the requirements for the degree of Doctor (PhD) in Science and Industrial Technologies from the Public University of Navarre.

https://doi.org/10.48035/Tesis/2454/45156

ACKNOWLEDGEMENTS

First, I would like to acknowledge my supervisors Dr. Jesús Álvarez Mozos and Dr. Miguel Ángel Campo Bescós for giving me the opportunity to develop this thesis over the last years. Thanks for all your time, help, support and guidance. I would also like to thank Dr. Luis Miguel Arregui for his collaboration in this thesis through the installation and setting up of the soil moisture probes in the experimental fields, and to Dr. María Gonzalez de Audícana for introducing me to the world of remote sensing during my master studies. I am also very grateful to all my colleagues from the THERRAE research group for all the good moments shared at the coffee breaks and lunchtime, and for all the assistance and advices in critical moments.

In addition, I am very grateful to the Institute for Earth Observation of EURAC research in Bolzano (Italy) for hosting me during my research stay during the beginning of 2022. It was a pleasure to meet you all and share nice experiences in Bolzano and the wonderful Dolomites. I want to specifically acknowledge Dr. Claudia Notarnicola for her time and expertise that helped me improving my work.

Likewise, thanks to all family and friends. Special thanks to my parents that were always helping me even during the difficult times and to Miguel, for his daily support no matter the distance. Thanks a lot!

Finally, this thesis has been possible thanks to a series of projects and a grand that are briefly listed in the following lines. This work was supported by the Spanish Ministry and Competitiveness and the European Regional Development Fund (MINECO/FEDER) through a project (CGL2016-75217-R) and a pre-doctoral grant (BES-2017-080560). It was also supported by the Spanish Ministry of Science and Innovation under Project PID2019-107386RB-I00/ AEI/10.13039/501100011033.

TABLE OF CONTENTS

LIST OF ACRONYMS

ACF	Autocorrelation function
AIEM	Advanced Integral Equation Model
ANN	Artificial Neural Networks
ASC	Ascending node
asl	Above sea level
BBCH	Biologische Bundesanstalt, Bimdessortenamt and Chemical Industry
BOA	Bottom of Atmosphere (BOA)
BON	Boletín Oficial de Navarra
CAP	Common Agricultural Policy
CCI	Climate Change Initiative
DEM	Digital Elevation Model
DESC	Descending node
ECV	Essential Climate Variable
Ens	Ensemble scheme
ESA	European Space Agency
EU	European Union
FAOSTAT	The Food and Agriculture Organization Corporate Statistical Database
Gamma-MAP	Gamma Maximum-A-Posteriori
GCOS	Global Climate Observing System
GEE	Google Earth Engine
GIS	Geographical Information System
GRD	Ground Range Detected
IEEE	Institute of Electrical and Electronics Engineers
IEM	Integral Equation Model
IQR	Interquartile Range
IW	Interferometric Wide
LAI	Leaf Area Index
LPIS	Land Parcel Information System
MIMICS	Michigan Microwave Canopy Scattering
MTBCD	Multitemporal Bayesian Change Detection
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near Infrared
OA	Overall Accuracy
PA	Producer's Accuracy
RBF	Radial Basis Function
reNDVI	Red-edge Normalized Difference Vegetation Index
RVI	Radar Vegetation Index
S1	Sentinel-1

S2	Sentinel-2
SAR	Synthetic Aperture Radar
SIGPAC	Sistema de Información Geográfica para la Política Agraria Común
SM	Soil Moisture
SNAP	Sentinel Application Platform
SRTM HGT	Shuttle Radar Topography Mission Height
STCD	Short Term Change Detection
SVM	Support Vector Machines
SVR	Support Vector Regression
SWIR	Short-wave Infrared
TIR	Thermal Infrared
TUWCD	TU Wien Change Detection
UA	User's Accuracy
UAV	Unmanned Aerial Vehicles
UAVSAR	Uninhabited Aerial Vehicle Synthetic Aperture Radar
UNFCCC	United Nations Framework Convention on Climate Change
VH	Vertical transmit, Horizontal receive polarization mode
VV	Vertical transmit, Vertical receive polarization mode
VWC	Vegetation Water Content
WATCOR	Wheat Attenuation Correction
WCM	Water Cloud Model

LIST OF SYMBOLS

α_{pp}	Fresnel reflection coefficient at HH or VV polarization (STCD method)
α_{min}, α_{max}	Boundary conditions of Fresnel reflection coefficient at VV polarization (STCD method)
σ^0	Sigma nought backscatter coefficient
σ^0_{can}	Total canopy backscatter (WCM method)
σ^0_{veg}	Backscatter from the vegetation (WCM method)
σ^0_{soil}	Backscatter from the soil (WCM method)
β^0	Beta nought backscatter coefficient
τ^2	Two-way attenuation factor (WCM method)
γ^0	Gamma nought backscatter coefficient
$\gamma^0_{(ref)}$	Gamma nought backscatter coefficient normalized at a reference incidence angle
$\gamma^0_{(ref)az\_corr}$	Gamma nought backscatter coefficient normalized at a reference incidence angle with azimuthal anisotropy correction
$\gamma^0_{i,IEM}$	Theoretical backscatter values calculated from the ground SM measurements (MTBCD metod)
γ^0_{CAN}	Gamma nought backscatter coefficient (Chapter 6)
γ^0_{SOIL}	WATCOR Gamma nought backscatter coefficient (Chapter 6)
θ_{loc}	Local Incidence angle
$\|\Delta\|$	Absolute backscatter difference (Evaluation of acquisition geometry)
$\Delta\gamma^0$	Difference between subsequent Sentinel-1 backscatter observations (WATCOR correlation analysis)
ΔSM	Difference between subsequent soil moisture measurements (WATCOR correlation analysis)
μ	Mean
σ	Standard deviation
A	Empirical parameter (WCM method)
B	Empirical parameter (WCM method)
C	Soil parameter (WCM method)
D	Soil parameter (WCM method)
d_{orbit}	Correction factor to compensate azimuthal effects
dB	Decibel
ε	Soil dielectric constant
ε_{min}	Minimum value of the uniform density function (MTBCD method)
F_k	Fitting index (WATCOR method)
JM_{ij}	Jeffries-Matusita distance
k	Number of change points to estimate (Change point analysis algorithm)

l_{corr} or l	Correlation length
min_size	Minimum number of observations between change points (Change point analysis algorithm)
MSE	Mean Squared Error
p	Order of the polynomial least-squared fit (Savitzky-Golay filter)
P_{prior}	A priori joint density function for ε (MTBCD method)
$P_{post}(\gamma_1^0, \gamma_2^0 \mid \varepsilon)$	Posterior density function (MTBCD method)
$P(\gamma_1^0, \gamma_2^0)$	Normalization factor (MTBCD method)
$P(R)$	Probability density (MTBCD method)
R	Pearson correlation coefficient
R	Random variable (MTBCD method)
R^2	Coefficient of determination
RMSE	Root Man Squared Error
s	Standard deviation of heights
SM_{diff}	Mean monthly differences between the maximum and minimum SM (STCD method)
SM_{est}	Estimated volumetric SM
SM_{obs}	Volumetric soil moisture recorded by the probes
SM_{sat}	Saturated soil moisture content (TUWCD method)
SM_{wp}	Wilting point (TUWCD method)
SSM	Surface Soil Moisture content (TUWCD method)
t, γ^0	Backscatter time series (WATCOR method)
t, γ_{s0}^0	Smoothed curve (WATCOR method)
$(t, \gamma^0)_{ref}$	Local mínimum points in the backscatter time series (WATCOR method)
ubRMSE	Unbiased root-mean-square error
V_1, V_2	Vegetation descriptors (WCM method)
w	Width of moving window (Savitzky-Golay filter)

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Crop monitoring is essential for different applications such as food security assurance, crop management, and the design and implementation of agricultural policies. Remote sensing provides information about biophysical properties of plants and soils and their spatial variability on large areas of the territory on a periodic basis. The launch of Sentinel-1 synthetic aperture radar (SAR) satellites in 2014 and 2016 made possible the acquisition of dense time series of images with good spatial resolution and temporal resolution even in cloud-covered areas. The main objective of this thesis is the evaluation of different methodologies for agricultural applications at the field scale using Sentinel-1 time series. First, a supervised crop classification methodology based on time signatures from Sentinel-1 images was proposed and implemented in a case study with 14 crop classes and a large dataset of agricultural fields. Secondly, the influence of acquisition geometry of Sentinel-1 images over wheat fields was assessed. The influence of the incidence angle on backscatter coefficients and terrain-flattened coefficients was evaluated, and an incidence angle normalization followed by an azimuthal anisotropy correction were applied to VV polarized time series. Thirdly, the backscatter attenuation produced by wheat canopy in VV polarization was evaluated, and a new wheat attenuation correction methodology named WATCOR was proposed. Finally, four different soil moisture (SM) estimation techniques based on Sentinel-1 time series were evaluated over wheat fields, proposing different methodological alternatives for their application at the field scale. Despite the complexity of estimating SM at the field scale solely with SAR data, acceptable estimations were obtained. The results of this thesis showed that the analysis and extraction of the information contained in SAR time series is useful for various agricultural applications, foreshadowing exciting future developments in this field.

RESUMEN

La monitorización de los cultivos es esencial para diferentes aplicaciones, como el aseguramiento de la seguridad alimentaria, la gestión de los cultivos y la implementación de políticas agrarias. La teledetección proporciona información acerca de las propiedades biofísicas de las plantas y los suelos, así como de la variabilidad espacial en amplias áreas del territorio de forma periódica. El lanzamiento de los satélites radar de apertura sintética (SAR) Sentinel-1 en 2014 y 2016 permitió la adquisición de series temporales densas de imágenes con buena resolución espacial y temporal incluso en zonas cubiertas de nubes. El principal objetivo de esta tesis es la evaluación de diferentes metodologías para aplicaciones agrícolas a escala de parcela usando series temporales Sentinel-1. En primer lugar, se propuso una metodología de clasificación de cultivos supervisada basada en las firmas temporales de Sentinel-1. Se implementó en un caso de estudio con 14 clases de cultivos y un dataset grande de parcelas agrícolas. En segundo lugar, se evaluó la influencia de la geometría de adquisición de las imágenes Sentinel-1 sobre parcelas de trigo. Se evaluó la influencia del ángulo de incidencia en la retrodispersión y los coeficientes de terrain-flattening, y se aplicó una normalización del ángulo de incidencia seguida de una corrección de la anisotropía azimutal en las series temporales de polarización VV. En tercer lugar, se evaluó la atenuación de la retrodispersión producida por las cubiertas de trigo en polarización VV, y un nuevo método de corrección del efecto de la atenuación llamado WATCOR fue propuesto. Finalmente, cuatro técnicas para estimar la humedad del suelo basadas en series temporales de Sentinel-1 en cultivo de trigo fueron evaluadas, proponiendo diferentes alternativas metodológicas para su aplicación a escala de parcela. A pesar de la complejidad de la estimación de la humedad del suelo a escala de parcela únicamente con datos SAR, se obtuvieron estimaciones aceptables. Los resultados de esta tesis demostraron que el análisis y la extracción de información contenida en series temporales SAR es útil para diferentes aplicaciones, augurando interesantes desarrollos futuros en este campo.

EXECUTIVE SUMMARY

Crop monitoring is important for several applications, such as food security assurance, crop management, and the design and implementation of agricultural policies. Remote sensing enables the acquisition of land cover information on a periodic basis, providing information on the biophysical properties of plants and soil, as well as their spatial variability. Radar sensors are active sensors that provide information on the geometric and dielectric characteristics of the surface, offering complementary information to optical sensors. One of their main advantages is that cloud cover does not affect the acquisition of images. The launch of the Sentinel-1 synthetic aperture radar (SAR) satellites in 2014 and 2016 made it possible to obtain dense time series of images with good spatial and temporal resolution free of charge. These time series provide an excellent opportunity for monitoring dynamic processes, such as vegetation growth or soil moisture changes at the agricultural field scale.

The broad objective of this thesis is to evaluate different methodologies for high spatial resolution agricultural applications based on Sentinel-1 time series. This has been achieved through several specific objectives: 1) the proposal and evaluation of a crop classification scheme based on the information contained in the time signatures of different crops. 2) The evaluation of the influence of acquisition geometry in wheat backscatter time series and its correction. 3) The characterization of the backscattering behavior of wheat in VV polarization and the proposal and evaluation of a new method that corrects its attenuation effect based on backscatter time series. 4) The evaluation of different multitemporal methods to estimate soil moisture in wheat fields.

First, a supervised crop classification methodology based on time signatures extracted from Sentinel-1 imagery was implemented in a case study with 14 crop classes and a large dataset of agricultural fields. The methodology was tested at the regional and provincial level, and the influence of field size in the classification accuracy was also evaluated. The results obtained varied depending on the input variables used. When considering the two polarizations (VV and VH) and their ratio (VH/VV) overall accuracies above 70% were achieved. Crops that showed particular time signatures achieved better results, such as barley, rice, corn and wheat. Field size had an influence on the accuracy of the results, with larger fields (>1 ha) achieving better scores than smaller fields (<0.5 ha). When considering the classification at the regional (sub-provincial) level, the obtained results improved. However, regions with a greater diversity of crops, different agricultural management practices, and higher number of fallow fields, obtained worse results. Due to its simplicity, the methodology could be easily implemented for Common Agricultural Policy (PAC) monitoring procedures.

Then, the influence of Sentinel-1 image acquisition geometry on the backscatter time series of wheat fields was evaluated. For this purpose, a dataset containing ~40,000 wheat fields for two agricultural years was used. The influence of the incidence angle on backscatter coefficients (σ^0) and terrain-flattened backscatter coefficients (γ^0) was evaluated. The terrain-flattening algorithm reduced the dependence of backscatter on the incidence angle, especially in VH polarization. Yet, in VV polarization some incidence angle influence remained. Interestingly, backscatter dependence on image geometry varied throughout the agricultural year, due to the phenological development of wheat plants, For reducing this influence, an incidence angle normalization technique was applied, followed by an azimuthal anisotropy correction. In summary, γ^0 allowed an acceptable combination of different relative orbits, although for applications requiring a complete independence of acquisition geometry, further corrections might be necessary in order to remove the orbital differences, especially in periods of bare soil or in winter.

Subsequently, it was observed that wheat canopy attenuated backscatter in VV polarization, defining a characteristic temporal pattern. For evaluating this attenuation a large dataset of wheat fields was analyzed, consisting of ~80000 wheat fields for four agricultural years. Sentinel-1 backscatter time series were obtained for these fields, proving that the attenuation pattern was consistent. A new wheat attenuation correction methodology (WATCOR) was proposed with the aim to correct this attenuation. The hypothesis of this method is that in the absence of attenuation, backscatter would follow a stable long-term trend with short-term variations caused by changes in soil moisture. The method was based on standard time series analysis techniques: Savitzky-Golay smoothing, changing point detection and lower envelope fitting, with no the need of external data (i.e., other vegetation descriptors like NDVI, etc.). The performance of WATCOR was assessed by measuring the correlation between backscatter and soil moisture measurements on six experimental wheat fields. The Water Cloud Model (WCM) vegetation correction was also applied as a reference method. The results showed that after applying WATCOR wheat attenuation was successfully removed and the highest correlation values were obtained between backscatter and soil moisture, improving significantly the correlations obtained with other options.

Finally, four different methodologies were evaluated for estimating soil moisture over wheat fields based on Sentinel-1 time series. Three change detection methods were selected: Short Term Change Detection (STCD), TU Wien Change Detection (TUWCD) and Multitemporal Bayesian Change Detection (MTBCD), and a Machine Learning technique: Suport Vector Regression (SVR). Different methodological issues were evaluated for the implementation of these techniques at the agricultural field

scale, including the calibration of some required parameters and the selection of the optimal methodological alternatives. Eight experimental wheat fields with available soil moisture measurements were used for validating the methods using different performance metrics. In particular, the use of canopy backscatter observations or vegetation corrected (WATCOR) backscatter values was compared. Furthermore, the influence of different factors such as the satellite pass, the type of management (rainfed or irrigated), the month of the year or the soil moisture content was also analyzed. The results were rather variable, with some experimental fields achieving successful performance metrics and some others rather poor ones. In general, it was observed that both TUWCD and MTBCD methods obtained better results when run with vegetation corrected backscatter time series. On the other hand, STCD and SVR produced similar results with and without vegetation correction. Both the month and the soil moisture content had an influence on the accuracy of the different methodologies. In most cases soil moisture was overestimated for dry conditions, and, to a lesser extent, it was underestimated for wet conditions.

The results obtained from this doctoral thesis demonstrate the usefulness of dense Sentinel-1 backscatter time series for agricultural applications at the field scale, especially in such an important crop as wheat. Therefore, it is expected that this work will encourage further research and in the longer term the operational use of these techniques by different stakeholders in the agricultural sector.

PUBLICATIONS

This doctoral thesis is presented as a collection of publications on the broad topic of investigating high spatial resolution agricultural applications based on Sentinel-1 time series. In particular, three articles have already been published and a fourth one is being considered for publication at the time of writing this thesis:

- Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2020. Crop Classification Based on Sentinel-1 Observations over Navarre Province, Spain. *Remote Sensing*, 12, 1-29. DOI: 10.3390/rs12020278.
  Bibliometric data: Impact factor: 5.349 (2021), Subject category: GEOSCIENCES, MULTIDISCIPLINARY, rank 30 out of 201 (Q1). 46 citations.

- Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2022. On the influence of acquisition geometry in backscatter time series over wheat. *International Journal of Applied Earth Observation and Geoinformation*, 106, 102671. DOI: 10.1016/j.jag.2021.102671.
  Bibliometric data: Impact factor: 7.672 (2021). Subject category: REMOTE SENSING, rank 5 out of 34 (Q1).

- Arias, M., Campo-Bescós, M.A., Arregui, L.M. González-Audícana, M. and Álvarez-Mozos J. 2022. A new methodology for wheat attenuation correction at C-band VV-polarized backscatter time series. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 4411314. DOI: 10.1109/TGRS.2022.3176144.
  Bibliometric data: Impact factor: 8.125 (2021). Subject category: REMOTE SENSING, rank 4 out of 34 (Q1).

- Arias, M., Notarnicola, C. Campo-Bescós, M.A., Arregui, L.M. and Álvarez-Mozos J. Evaluation of soil moisture estimation techniques based on Sentinel-1 observations over wheat fields. Submitted to *Science of the Total Environment* on November 2022.
  Bibliometric data: Impact factor: 10.753 (2021). Subject category: ENVIRONMENTAL SCIENCE, rank 15 out of 325 (Q1).

In addition to these articles, which are a result of this doctoral thesis, five presentations at national and international conferences were carried out:

- Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2018. Crop Type Mapping Based on Sentinel-1 Time Series. *International Geoscience and Remote Sensing Symposium (IGARSS)* 2018, Valencia, 22-27 July. Oral communication presented by Jesús Álvarez Mozos.

- Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2019. Crop classification using Sentinel-1 and Sentinel-2 time series. *ESA Living Planet Symposium.* Milan, 13-17 May 2019. Poster presented by María Arias Cuenca.

- Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2019. Análisis de sensibilidad de la técnica multitemporal de estimación de humedad superficial del suelo Alpha approximation. *XVII Congeso de la Asociación Española de Teledetección: hacia una visión global del cambio climático*. Valladolid, 25-27 September 2019. Poster presented by María Arias Cuenca.

- Arias, M., Campo-Bescós, M.A., Arregui L.M., González-Audícana, M. and Álvarez-Mozos J. 2022. Correction of wheat attenuation in Sentinel-1 VV-polarized backscatter time series. *ESA Living Planet Symposium,* Bonn. 23-27 May 2022. Poster presented by María Arias Cuenca.

- Arias, M., Campo-Bescós, M.A., Arregui L.M. and Álvarez-Mozos J. 2022. Evaluación de un nuevo método para la corrección de la atenuación de retrodispersión producida por trigo en series temporales de Sentinel-1. *XIX Congreso de la Asociación Española de Teledetección: Teledetección para una Agricultura Sostenible en la era del Big Data.* Pamplona, 29 June – 1 July 2022. Oral communication presented by María Arias Cuenca.

Following the regulations of the Doctoral School of Navarre (EDONA) from the Public University of Navarre (UPNA), the document is structured in this way: first, a chapter of introduction and a chapter of materials and methods are presented. Then, the articles are included as they were published. Finally, it ends with some general conclusions.

CHAPTER 1

INTRODUCTION

1.1. General introduction

The continuous increase in food demand associated with the human population growth, along with the scarcity of fertile agricultural land and water resources and the impacts of climate change are the main challenges that the agricultural sector has to face to guarantee future food security (FAO, 2018; Orynbaikyzy et al., 2019). This thesis focuses in wheat, as it is one of the most important staple food in the world, with more than 200 million ha cultivated and 780 million tons produced per year (FAO, 2021).

In this context, the monitoring of crops and agricultural activities is essential (Atzberger, 2013), because as the saying goes ''if you can't measure it, you can't manage it'. Conventional land monitoring methods are based in fieldwork, being expensive and time-consuming (Chuvieco, 2019). Remote sensing techniques derive information from the radiation reflected or emitted by earth surfaces measured from sensors in distant platforms, such as unmanned aerial vehicles (UAV), aircrafts or satellites (Mulla, 2013). At present, they represent an efficient and effective alternative for land monitoring, providing information about the biophysical characteristics of plants and soil, and their spatial variability with a given periodicity.

Landsat 1, the first satellite for natural resource monitoring, was launched in 1972, initiating the space remote sensing era. In the following years, the first agricultural remote sensing works based on optical imagery were carried out (Jewell, 1989; Misra and Wheeler, 1978; Rouse et al., 1974; Wheeler and Misra, 1980). Generally, optical sensors capture the energy from the sun that is reflected by land surfaces in the wavelengths of the visible, near infrared (NIR) and short-wave infrared (SWIR), and can be used to estimate different parameters such as the leaf area index (LAI), biomass, chlorophyll content or vegetation water content (VWC) (LIU et al., 2019). Despite the popularity of optical imagery for agricultural applications, it is affected by weather conditions. Hence, monitoring in the optical domain cannot be implemented in many parts of the world in an operational way due to persistent cloud cover (Whitcraft et al., 2015) that limits the number of available observations during the agricultural season.

Synthetic Aperture Radars (SAR) are active sensors that operate in the microwave region of the electromagnetic spectrum (cm wavelengths). They emit microwave pulses towards the Earth surface and then measure the echoes that return back to the sensor after interacting with the land surface elements. Depending on the characteristics of the observed surface and the configuration of the sensor, the pulse undergoes different scattering and reflection processes (Ulaby et al., 1982). The ratio between the power of the echo received back by the sensor and the power of the pulse initially emitted represents the backscatter coefficient. The main sensor configuration parameters that govern the backscatter response are the frequency or wavelength used, the incidence

angle and the polarization emitted and received by the sensor. Observation frequency determines the size of the scattering elements that contribute to backscatter (McNairn and Brisco, 2004). For instance, in densely vegetated targets, shorter wavelengths (X-band, λ= 2.5-4 cm) interact with small surface elements, like leaves and fruits; medium wavelengths (C-band, λ= 4-8 cm) are more sensitive to small branches or stems; while longer wavelengths (L-band, λ= 15-30 cm) interact with larger branches and trunks. The longer the wavelength the higher the capacity of the radar pulse to traverse the vegetation canopy and reach the soil, providing information related to soil moisture and soil surface roughness. The incidence angle is the angle between the direction of propagation of the radar pulse and the ground normal. It also influences the capability of the pulse to penetrate the vegetation canopy. This capability is higher with smaller incidence angles. The polarization is the orientation of the electric field of the pulse emitted and received by the sensor. In most Earth observation radar systems, polarization can be either vertical (V) or horizontal (H). Therefore, a sensor can receive and emit the pulse in the same orientation (co-polarization: VV or HH) or in different orientations (cross-polarization: VH or HV). Co-polarized backscatter is normally more sensitive to the soil response and double-bounce effects, while cross-polarized backscatter provides information about volume scattering, usually related to vegetation density or biomass.

One of the main advantages of SAR sensors is that they can operate in cloud covered conditions and backscatter observations are not affected by atmospheric components that absorb or reflect the radiation at visible or infrared frequencies. In addition, they can operate independently of sunlight conditions. Furthermore, SAR sensors offer suitable spatial resolution for agricultural monitoring at field scale (<30 m) and thus, they can be an alternative to optical imagery for agricultural applications. The information provided by these sensors is related to the geometric characteristic of the surfaces (shape, size and orientation of scatterers) and their dielectrical properties (soil moisture content and salinity). In this manner, they provide complementary information to optical sensors.

The first operational SAR satellites were launched in the decade of the nineties: ERS-1, ERS-2 and RADARSAT-1 (C-band), and JERS (L-band). In the 2000s, more satellite missions were launched: Envisat/ASAR and RADARSAT-2 (C-band), ALOS/PALSAR (L-band) and Cosmo-SkyMed and TerraSAR-X (X-band). During the last decade (2010) more satellites were launched: Risat-1, RCM, Sentinel-1A and Sentinel-1B (C-band), ALOS-2/PALSAR and SAOCOM-1a,b (L-band), NovaSAR-1 (S-band) and TanDEM-X, KOMPSAT-5, PAZ, ICEYE and Capella (X-band). For the 2020's decade many new SAR sensors will be launched to add-on and strengthen existing constellations (e.g., Sentinel-1 and ALOS/PALSAR), but also new innovative missions are scheduled for

launch, such as NISAR (dual frequency L- and S-band), Biomass (P-band), EOS SAR (dual frequency X- and S-band) or StriX-1 (X-band).

With the beginning of the Sentinel-1 mission in 2014, worldwide routine C-band observations became freely available (Berger et al., 2012). The Sentinel-1 mission is composed of two twin satellites (Sentinel-1A and Sentinel-1B) with a nominal temporal resolution of 6 days and a spatial resolution of 10-20m. Although Sentinel-1B mission ended prematurely on December 2021, due to a technical issue, Sentinel-1C is scheduled for launch in the first half of 2023 to ensure data continuity. Sentinel-1 is part of the Copernicus Programme, the European Union's Earth Observation programme. This program has been a turning point in the field of Earth observation, with routine observations potentially useful for many applications due to its excellent compromise between spatial resolution and revisit time. Specifically, the use of dense time series from satellite imagery allows monitoring the dynamic processes like vegetation growth or changes in soil moisture over time at field scale.

The combination of multiple orbits with different acquisition geometries (incidence and azimuth angles) that observe a specific zone could reduce the revisit time to less than two days in many European regions (Weiß et al., 2021). However, an assessment of the consistency of backscatter observations obtained with different geometries is necessary when combining data from different orbits (Gauthier et al., 1998). Indeed, backscatter depends on the incidence angle (Ulaby et al., 1982), although the magnitude of these variations also depends on the characteristics and scattering mechanisms of the observed surface (Ardila et al., 2010). Surfaces where specular scattering dominates are more sensitive to variations in the incidence angle (Skriver et al., 1999a), whereas rough surfaces and vegetated areas have a lower backscatter-incidence angle dependence due to the predominance of volume scattering. The azimuth angle is the angle between the radar beam projected on the surface and a reference direction that is usually the North (Schaufler et al., 2018). In areas with predefined directional structures where surface scattering dominates (i.e., tillage rows), the azimuthal effects might also be more noticeable. Furthermore, in areas where topography is rough, the incidence and azimuth angles have a stronger effect in backscatter. Regarding agricultural surfaces, the scattering mechanisms governing backscatter change during the growing cycle of crops and therefore, the influence of incidence and azimuth angles in backscatter might also change during the year,

C-band backscatter is sensitive to the phenological cycle of different crops (Mattia et al., 2003; Skriver et al., 1999b; Veloso et al., 2017). Therefore, dense time series have a high potential for tracking phenological changes, with different authors proving their utility for crop monitoring and for the identification of different phenological stages (Mercier et al., 2020; Nasrallah et al., 2019; Schlund and Erasmi, 2020; Song and Wang,

2019). The knowledge of the start and end of the phenological stages is important for farmers and decision makers, as it influences different agricultural practices such as irrigation, fertilization or pesticides application, and might also provide useful information for making yield predictions (Nasrallah et al., 2019). Furthermore, data on crop phenology and development might provide interesting information for crop classification (Bargiel, 2017), which is an important agricultural application of remote sensing, both at a local and global scale (Kussul et al., 2016). Crop maps and the information extracted from them are extremely valuable for the design, implementation and monitoring of agricultural policies, but also for forecasting yields and assuring food security (Schmedtmann et al., 2015; Van Tricht et al., 2018). In the case of SAR data, it is crucial to understand the temporal variations of backscatter in different crops before the development of operational monitoring methodologies (Veloso et al., 2017). With the use of time series, temporal models for individual crops can be built and later used in classification procedures (Whelen and Siqueira, 2017). The concept of these temporal models is that crops have physical differences during their growing season due to their phenological development, and consequently, lead to different backscatter time series. These models called temporal signatures represent the typical backscatter response of a crop over time, being representative of each specific crop.

Another important variable that is essential in agriculture is soil moisture (SM). It is crucial for crop growth because it determines the amount of water available for plants, but also influences other aspects of the agricultural activity, such the occurrence of plagues and diseases (Vereecken et al., 2010) or agricultural trafficability (Carranza et al., 2019). Microwave sensors are sensitive to the dielectrical properties of surfaces, and therefore, they are suitable for SM estimation. However, SAR observations, are also affected by other variables like surface roughness (Verhoest et al., 2008) or the characteristics of vegetation (Bindlish and Barros, 2001), and thus, the estimation of the actual SM content of a particular field using SAR data is complex. There are several models that have been developed over the years for SM estimation in bare soils (Baghdadi et al., 2016; Chen et al., 2015; Dubois et al., 1995; Fung, 1994; Oh, 2004; Oh et al., 1992; Shi et al., 1997). Their inversion allows calculating SM from backscatter observations if the other variables included in the model are known (typically surface roughness parameters). In the case of agricultural canopies, it is necessary to couple a model that simulates vegetation backscatter (Attema and Ulaby, 1978; Brown et al., 1994; Ulaby et al., 1990). These models require a large number of crop variables or a specific parameterization for each crop and location. Therefore, the estimation of SM from SAR observations is typically an ill-posed problem with too many unknown variables. To circumvent this, methods based on the combination of images acquired with different frequencies or incidence angles were developed (Jagdhuber et al., 2013;

Notarnicola and Posa, 2007) or also based on multi-polarization or polarimetric data. Nevertheless, the operational application of these alternatives at the field scale is limited due to the unavailability of the required input data in a routine basis. Machine learning algorithms have been also applied in this domain (Baghdadi et al., 2012; Liu et al., 2021; Pasolli et al., 2011), although they require field data for the training phase and might not be easily generalizable and applicable to other sites. However, dense SAR backscatter time series along with change detection techniques might provide a viable path for developing operational SAR based SM estimation methods (Balenzano et al., 2013; Wagner et al., 1999). The hypothesis of these techniques is that if time series are dense enough, the differences between consecutive backscatter observation are caused only by SM variations, because the other variables affecting backscatter change at a slower rate.

This introduction evidences that Sentinel-1 backscatter time series constitute an excellent data source with potential applications for agriculture, which need to be explored in detail through research. This is the rationale of this PhD thesis.

1.2. Objectives

As mentioned, the launch of the Sentinel-1 mission opened a great opportunity for developing agricultural applications at the field scale using C-band backscatter imagery. The general objective of this thesis is to propose, analyze and assess different data-driven methodologies for agricultural applications at the field scale based on Sentinel-1 SAR backscatter time series. The first considers different crops, while the second part focuses specifically on winter wheat. This general aim was achieved by the compliance of the following specific objectives:

- The proposal and evaluation of a crop classification scheme based on the information contained in the bckscatter temporal signatures of each crop, in a case study with a large number of classes and high heterogeneity of agro-climatic conditions and field sizes.
- The evaluation of the influence of acquisition geometry on wheat backscatter time series, and the proposal of a strategy for its correction.
- The characterization of the backscattering behavior of wheat in VV polarization, and the development of a new methodology to correct the attenuation effect of wheat with no need of external information.
- The implementation and assessment of different methodologies for the estimation of SM in wheat fields, based on Sentinel-1 time series.

1.3. Structure of the thesis

The thesis has been written as a compendium of research papers that were published or submitted for publication in international scientific journals. Consequently, since each work was intended to be read independently, there might be some overlap between them, particularly in the introduction and materials and methods sections.

1) The first chapter consists of a general introduction and state of the art where the objectives and justification of the thesis is presented.

2) The second chapter describes the imagery, the study area, the ground measurements and methodology used in the thesis.

The following chapters are the published (or submitted) scientific papers that are based on the particular objectives described in the previous section.

3) The third chapter proposes and evaluates a supervised crop classification algorithm based on backscatter temporal signatures, in a case study with a large number of classes, agro-climatic conditions and field sizes. Specifically, the separability and temporal evolution of crops is analyzed, and the influence on the different crop growing conditions and field sizes on the classification results are evaluated.

4) The fourth chapter evaluates the influence of acquisition geometry (incidence and azimuth angles) in Sentinel-1 backscatter time series over wheat canopies, proposing a strategy for its correction.

5) The fifth chapter describes the backscatter behavior of wheat at VV-polarization during its growing cycle, proposing a new methodology for wheat attenuation correction (WATCOR) based on the time series information itself, Then, the method is compared with other vegetation correction methodologies by measuring the correlation between the corrected backscatter and SM ground measurements in some experimental wheat fields.

6) The sixth chapter evaluates different SM estimation techniques in wheat fields based on Sentinel-1 time series, comparing the results obtained with the different techniques, and also evaluating the eventual benefits of vegetation correction with WATCOR before the SM estimation techniques are applied.

CHAPTER 2

MATERIALS AND METHODS

2.1. Satellite images

Sentinel-1 image time series were the base of this thesis. Yet, additional images obtained with other sensors were also used in chapters 4, 5 and 6.

2.1.1. Sentinel-1

Sentinel-1 satellites carry a SAR sensor that operates at C-band (5.4 GHz). All images used in this thesis were acquired in the Interferometric Wide (IW) swath mode and downloaded as Level-1 Ground Range Detected (GRD) products. The IW mode is the main acquisition mode over land. It uses a dual-polarization configuration VV-VH, and covers a 250 km swath at 5 m by 20 m spatial resolution. The Level-1 GRD products are focused SAR data that were multi-looked (5x1) and projected to ground range using the Earth ellipsoid model WGS84, obtaining a spatial resolution of ~20m, although products are distributed with a 10m pixel size. The images belonged to three relative orbits, that covered the study region (the province of Navarre, Spain): 103ASC, 8DESC and 81DESC. The acquisition times for each orbit were 18:03 UTC, 6:08 UTC and 6:17 UTC respectively.

All the images were processed using an automated pipeline in SNAP Graph Processing Toolbox, obtaining terrain-flattened γ^0 backscatter coefficient values in dB units and a pixel size of 20 m. A simpler pipeline was also applied in chapter 4, producing σ^0 backscatter values, for comparison with γ^0. The details about the pipelines can be found in each chapter. The main characteristics regarding acquisition time and number of images are shown in Table 2.1.

Table 2.1. Sentinel-1 images used in each Chapter.

Paper	Backscatter coefficient	Polarization channels	Period	Number of images	Temporal resolution
Arias et al. (2020) Chapter 3	γ^0	VH, VV, VH/VV	01/09/2015 – 31/08/2016	110	12 days
Arias et al. (2022a) Chapter 4	σ^0, γ^0	VH, VV	01/09/2016-31/08/2018	341	6 days
Arias et al. (2022b) Chapter 5	γ^0	VV, VH/VV	01/09/2015 – 31/08/2020	732	6 days, 12 days
Arias et al. (2022c) Chapter 6	γ^0	VV, VH, VH/VV	01/09/2015 – 31/08/2021	911	6 days, 12 days

2.1.2. Sentinel-2

Sentinel-2 mission also belongs to the Copernicus Programme and it comprises two identical (Sentinel-2A and Sentinel-2B) polar-orbiting satellites that acquire multispectral imagery. The revisit time of each satellite is 10 days (5 days combining A

and B) and the spatial resolution varies between 10 m, 20 m and 60 m depending on the spectral band. In this thesis, Sentinel-2 Level-2A Bottom of Atmosphere (BOA) products were used. Scenes were processed using Google Earth Engine that was used to discard cloud-contaminated scenes, resample images, calculate different spectral indices and obtain statistics for agricultural fields and regions of interest (Table 2.2.).

Table 2.2. Sentinel-2 images used in each Chapter.

Paper	Bands	Vegetation index	Period	Number of images
Arias et al. (2022a) Chapter 4	B5, B6	Red-edge NDVI (reNDVI)	01/09/2017-31/08/2018	14
Arias et al. (2022b) Chapter 5	B4, B8, B11	NDVI, NDWI	01/09/2017 – 31/08/2020	89
Arias et al. (2022c) Chapter 6	B4, B8	NDVI	01/09/2017 – 31/08/2021	120

2.1.3. ESA CCI Soil Moisture Product

In 2012, the European Space Agency (ESA) launched a global satellite-observed soil moisture dataset belonging to the Climate Change Initiative (CCI) program. This product called ESA CCI SM combines different single-sensor active and passive microwave soil moisture products, resulting in products: a merged active, a merged passive and a combination of active and passive microwave product (Dorigo et al., 2017). The product, with daily temporal resolution, presents a spatial resolution of 0.25^0. In this thesis, the combined product was used as ancillary information for one of the soil moisture estimation techniques in chapter 6. Soil moisture time series from 1978 until 2019 were downloaded from the following website: https://esa-soilmoisture-cci.org. In total, 15036 observations were obtained.

2.2. Study sites

In this thesis, two main study sites were considered. The first study site was the largest, and it comprised all the agricultural lands of the province of Navarre (this was used in chapters 3, 4, 5 and 6). Navarre is located in the northern central part of Spain, and it covers an area of 10391 km^2, out of which 1/3 is agricultural land. Due to its climatic and landscape diversity, Navarre is divided in seven agricultural regions (Figure 2.1). The second study site was much smaller and consisted of eight experimental fields cultivated with wheat, where ground soil moisture measurements were recorded (chapters 5 and 6). Figure 2.1. shows the location of the study sites.

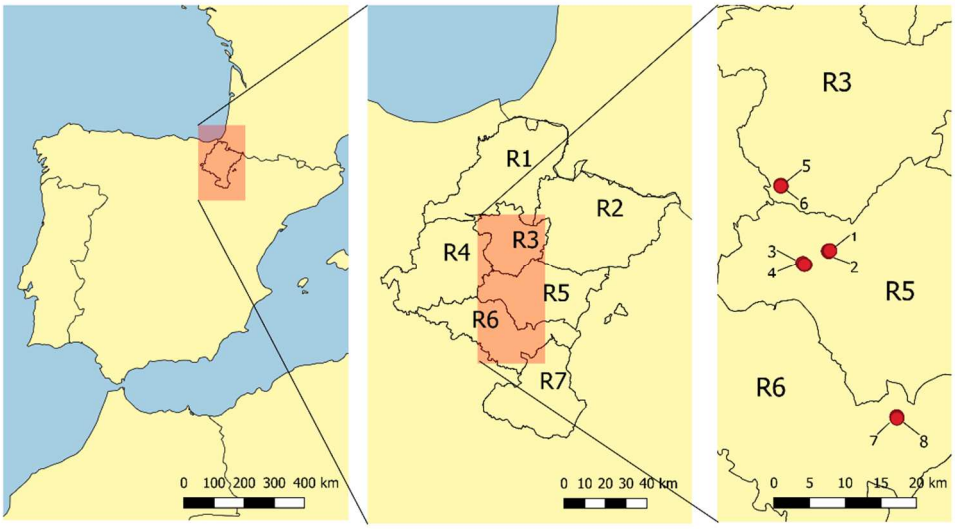

Figure 2.1. Location of the province of Navarre (left), the agricultural regions of Navarre (centre), and the experimental wheat fields (right).

2.2.1. Agricultural regions of Navarre

The province of Navarre, with a relatively small area (10391 km^2), is characterized by a variety of landscapes and climatological conditions. The Northern area, mainly covered by grasslands and woodlands, is mountainous and presents a humid climate. On the contrary, the Southern area is formed by the plains of the Ebro river basin and presents a drier climate. The proportion of arable land is higher in the Southern part, where both irrigated and rainfed agriculture are present. The area between these two zones is a transition zone with mixed characteristics. Due to this diversity, Navarre is divided into seven agricultural regions for administrative and management purposes.

The agricultural fields used in this thesis were extracted from a large geo-database containing information about the declarations and inspections for the EU Common Agricultural Policy (CAP) in Navarre. This database was provided by the Agriculture Department of the Government of Navarre as an anonymized version. The dataset consisted of GIS polygons representing the field boundaries of the crops stated by farmers in their CAP declarations, and the field boundaries of the parcels inspected by the Government technicians (Table 2.3). The database is divided in agricultural years. Each agricultural year starts the 1st September of the previous year and finishes the 31st August of the agricultural year.

Table 2.3. Characteristics about the CAP databases used

Paper	Type of CAP data	Agricultural Regions	Crops	Agricultural years	Number of fields
Arias et al. (2020) Chapter 3	Inspections and declarations	1-7	14 classes	2016	121093
Arias et al. (2022a) Chapter 4	Declarations	1- 7	Wheat	2017, 2018	39124
Arias et al. (2022b) Chapter 5	Declarations	1- 7	Wheat	2016-2019	81434
Arias et al. (2022c) Chapter 6	Declarations	3, 5, 6	Wheat	2016-2019	43595

2.2.2. Experimental wheat fields

The winter wheat experimental fields were located in regions 3, 5 and 6 (Figure 2.1). The study period covered four agricultural campaigns: 2018, 2019, 2020 and 2021. In each agricultural campaign, two fields were monitored: one rainfed and the other irrigated. Detailed information about their size, soil texture, sowing and harvesting dates can be found in chapter 5 (Section 5.4.1) and chapter 6 (Section 6.2.1).

Ground soil moisture measurements were acquired for training and validating remote sensing estimates. On each field, from three to five capacitance SM probes were installed in winter and removed before harvest. Sentek-multi probes were used (Figure 2.2.), which record the volumetric SM every 30 minutes at 6 different depths, from the surface to 60 cm deep.

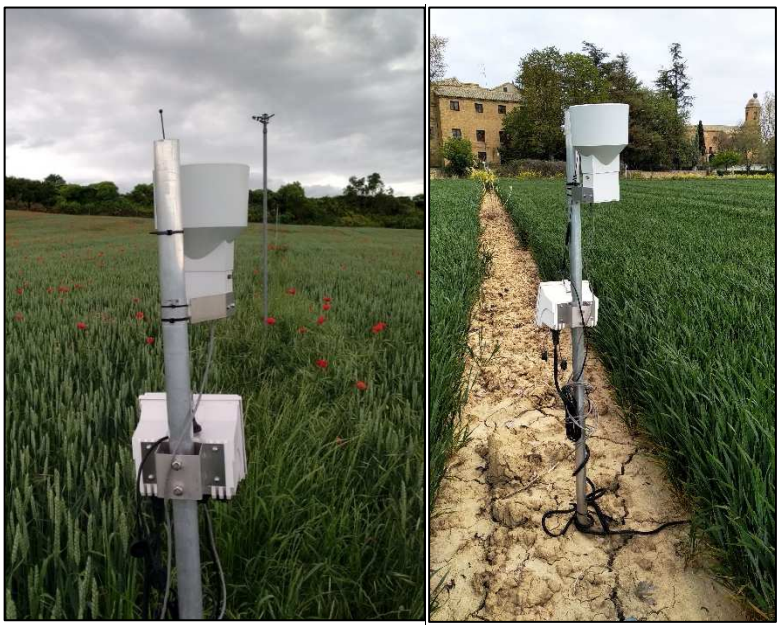

Figure 2.2. Sentek-multi soil moisture probes in field 5 (left) and field 7 (right).

2.3. Methodology

Table 2.4 shows an overview of the methodology, variables and performance metrics used in this thesis. Detailed information is explained in the materials and methods section of each chapter.

Table 2.4. Summary of the methodology used in each Chapter.

Chapter	Target variable	Input variables	Technique	Performance metrics
3	Crop class	γ^0_{VH}, γ^0_{VV}, $\gamma^0_{VH}/\gamma^0_{VV}$	Backscatter time series comparison with crop temporal signatures	JM Distance, OA, PA, UA, F1-score
4	Backscatter dependence on acquisition geometry	σ^0_{VH}, σ^0_{VV} γ^0_{VH}, γ^0_{VV}, θ_{loc}, reNDVI	Terrain flattening, incidence angle normalization, azimuthal anisotropy correction	Slope (dB/º), Absolute differences between orbit passes (dB), R
5	γ^0_{VV} wheat backscatter time series without attenuation effect	γ^0_{VV}, NDVI, NDWI, $\gamma^0_{VH}/\gamma^0_{VV}$	Wheat attenuation correction (WATCOR) based on time series analysis	Visual analysis, R
6	Soil moisture	γ^0_{VV}, γ^0_{VV} (WATCOR), NDVI, ESA CCI Soil Moisture Product	Short Term Change Detection, TU Wien Change Detection, Multitemporal Bayesian change detection, Support Vector Regression	RMSE (m^3/m^3), BIAS(m^3/m^3), R, ubRMSE(m^3/m^3)

CHAPTER 3

CROP CLASSIFICATION BASED ON SENTINEL-1 TIME SERIES

Published in: Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2020. Crop Classification Based on Sentinel-1 Observations over Navarre Province, Spain. Remote Sensing 2020, 12, 1-29. DOI: 10.3390/rs12020278.

Supplementary materials are available at https://www.mdpi.com/2072-4292/12/2/278

Abstract

Crop classification provides relevant information for crop management, food security assurance and agricultural policy design. The availability of Sentinel-1 image time series, with a very short revisit time and high spatial resolution, has great potential for crop classification in regions with pervasive cloud cover. Dense image time series enable the implementation of supervised crop classification schemes based on the comparison of the time series of the element to classify with the temporal signatures of the considered crops. The main objective of this study is to investigate the performance of a supervised crop classification approach based on crop temporal signatures obtained from Sentinel-1 time series in a challenging case study with a large number of crops and a high heterogeneity in terms of agro-climatic conditions and field sizes. The case study considered a large dataset on the Spanish province of Navarre in the framework of the verification of Common Agricultural Policy (CAP) subsidies. Navarre presents a large agro-climatic diversity with persistent cloud cover areas, and therefore, the technique was implemented both at the provincial and regional scale. In total, 14 crop classes were considered, including different winter crops, summer crops, permanent crops and fallow. Classification results varied depending on the set of input features considered, obtaining Overall Accuracies higher than 70% when the three (VH, VV and VH/VV) channels were used as the input. Crops exhibiting singularities in their temporal signatures were more easily identified, with barley, rice, corn and wheat achieving F1-scores above 75%. The size of fields severely affected classification performance, with ~14% better classification performance for larger fields (>1ha) in comparison to smaller fields (<0.5 ha). Results improved when agro-climatic diversity was taken into account through regional stratification. It was observed that regions with a higher diversity of crop types, management techniques and a larger proportion of fallow fields obtained lower accuracies. The approach is simple and can be easily implemented operationally to aid CAP inspection procedures or for other purposes.

Keywords: crop classification; Sentinel-1; SAR; time series; Common Agricultural Policy.

3.1. Introduction

Crop classification is one of the most important agricultural applications of remote sensing (Atzberger, 2013). Knowing what crops are present in the fields is very useful both at a local and global scale (Kussul et al., 2016). For instance, this information is valuable for the design and implementation of agricultural policies (Schmedtmann and Campagnolo, 2015), as well as for crop management and food security assurance (Van Tricht et al., 2018). Crop classification is also a pre-requisite for implementing other remote-sensing-based applications at the field scale (e.g., biomass and yield estimation, anomaly detection, etc.). Satellite Earth observations provide information about biophysical properties of the Earth's surface and their spatial variability with a given revisit time. This constitutes a very rich information source that can be used for identifying the crop types being cultivated and also, for monitoring them along their growing cycle (McNairn et al., 2002).

European farmers benefit from the Common Agricultural Policy (CAP) support of the European Union, which started back in 1962. CAP establishes different financing instruments including area-based payments and cross-compliance requirements (Regulation (EU) No 1306/2013). Competent authorities might check whether farmers' declarations are correct (i.e., they are actually growing the crops they declare on each field and follow cross-compliance requirements). Traditionally, this has been done through field inspections on a sample of fields, constituting an expensive, inefficient and incomplete check system. On June 2018, the European Commission drafted the new modification of the CAP (Regulation (EU) 2018/746) regarding applications, payment claims and checks to be fully implemented in 2020. These modifications recommend the establishment of procedures to check and track all eligibility criteria using Copernicus Sentinels data or similar data. Therefore, several initiatives and research efforts are being conducted at present to fulfill this aim, including EU-funded projects like Sen2-Agri (Defourny et al., 2019), RECAP (Sitokonstantinou et al., 2018) or SEN4CAP (Koetz et al., 2019).

Indeed, the Copernicus program has been a game changer in this field of research, with frequent observations potentially useful for crop classification (both in the microwave (Sentinel-1, S1) and optical (Sentinel-2, S2) domain), which are acquired systematically over the world and freely distributed (Berger et al., 2012). Their excellent compromise between spatial and spectral resolution, and above all, their enhanced revisit time, through the use of constellations of twin satellites (S1A – S1B and S2A – S2B), make them a unique solution for monitoring dynamic processes like vegetation growth at the scale of agricultural fields.

Until the Sentinels were launched, typical approaches for crop classification were mostly based on the evaluation of the spectral signature of crops obtained from multispectral imagery on key moments of the phenological cycle and their statistical (di)similarity (Chuvieco and Huete, 2009). However, early studies (Misra and Wheeler, 1978) already recalled that this approach might have limitations after observing that different fields of the same crop might have different spectral responses due to slight differences in the agricultural management calendar from field to field. In addition, different crops might expose a similar spectral response on a given date, so an examination of the multi-temporal evolution of the spectral response of a field was hypothesized as a potentially relevant information for crop recognition (Misra and Wheeler, 1978). Yet the unavailability of multi-temporal datasets with the sufficient revisit time (i.e., 16 days for Landsat) precluded the further development of this type of method until sensors like MODIS became available (Lobell and Asner, 2004; Wardlow et al., 2007). With the availability of S2 data, similar approaches but with a higher spatial resolution have been developed (Immitzer et al., 2016; Petitjean et al., 2012; Sitokonstantinou et al., 2018).

In large parts of the world, remote sensing in the optical domain has been difficult to implement operationally due to persistent cloud cover limiting the number of viable observations along the agricultural season (Whitcraft et al., 2015). Synthetic Aperture Radar (SAR) sensors can be an interesting alternative for those regions since they can operate in cloud-covered and poorly illuminated conditions. In addition, SAR data provide complementary information to optical sensors, reflecting geometric and dielectric properties of vegetation and soil, which have been proven relevant for crop classification (McNairn et al., 2009b; Steele-Dunne et al., 2017). The first SAR-based crop classification experiences used a few single polarization scenes and generally achieved limited results (McNairn et al., 2002), but soon after, several studies demonstrated that adding dimensionality through multi-frequency, multi-polarization or multi-temporal SAR observations increased the accuracy of crop maps (Hoekman and Vissers, 2003; Skriver et al., 2011). Observation frequency determines the size of the scattering elements that contribute to backscatter (McNairn and Brisco, 2004), with short wavelengths (X-band) interacting with smaller canopy constituents (e.g., leaves and fruits) than medium (C-band, e.g., small branches and stems) or longer wavelengths (L-band, e.g., larger branches and trunks). The latter might also be more sensitive to soil surface roughness and moisture conditions below the canopy. Regarding polarization, quad-polarized data provided polarimetric features which are potentially interesting for classification and other agricultural applications (Lee et al., 2001; McNairn and Brisco, 2004); however, dual-polarized configurations might be more efficient for multi-temporal large-area mapping and still provide interesting information for crop monitoring (Veloso et al., 2017) and classification (Larrañaga and Álvarez-Mozos, 2016;

Skriver, 2012). On the other hand, multi-temporal observations of single or dual-polarized backscatter have demonstrated a clear benefit on the accuracy of the resulting classifications (Skriver, 2012), since morphological differences in some particular stages of the phenological development might be key for differentiating pairs of crops that are otherwise quite similar the rest of the year (Steele-Dunne et al., 2017).

Given the actual availability of SAR constellation missions with an extremely short revisit time and high spatial resolutions, classification methods that explicitly exploit time series information have recently been developed with promising results. In effect, supervised classification schemes were developed based solely on time series information, where training data were used to obtain model time series for each class (i.e., temporal signatures) that were later compared with the elements to classify using some type of error metric. This way, Whelen and Siqueira (Whelen and Siqueira, 2017) classified alfalfa, corn and wheat fields over a 10 km x 23 km area using L-band UAVSAR observations, and obtained accuracy values higher than 75% when HV backscatter, entropy or alpha time series were used. Later, using C-band Sentinel-1 data, this approach was refined with a case study over a portion of central North Dakota considering four crops (corn, soybean, wheat and grass/pasture) and obtaining an accuracy of 85% (Whelen and Siqueira, 2018). In a similar fashion, Xu et al. (Xu et al., 2019) classified two corn-dominated sites using Sentinel-1 time series, and obtained accuracies ~90%. In these studies, the error metrics used to compare the temporal signatures with the time series of the elements to classify differed, but only slight differences were observed in the obtained accuracies (Xu et al., 2019). The performance of this classification approach was comparable to results obtained with machine learning algorithms. Yet, the simplicity and robustness of the approach makes it transparent, portable from season to season and resilient to crop growth anomalies. However, further studies were recommended to explore the performance of this approach on regions with a higher crop diversity (Whelen and Siqueira, 2018) and more heterogeneous conditions.

Indeed, the number of crop classes considered strongly impacts the accuracy of the results obtained (Skriver, 2012). Typically, published crop classification studies based on SAR data considered six different crop classes or less (McNairn et al., 2009b; Whelen and Siqueira, 2017; Xu et al., 2019), although some studies had a higher number—but this was normally below ten (Hütt et al., 2016; Larrañaga and Álvarez-Mozos, 2016) or fifteen (Bargiel, 2017; Skriver, 2012). Furthermore, some studies did not solely address crop classification, as they considered both crop categories and other land covers such as forests, water surfaces or urban (Hütt et al., 2016). Regarding the number of fields used for training and validation, the published studies differ greatly, although in many cases, they were below 100 (Whelen and Siqueira, 2017; Xu et al., 2019) or 1,000

(Bargiel, 2017; Skriver, 2012). Few published works considered near-operational applications with 10,000–100,000 fields or more and these used optical imagery (Defourny et al., 2019; Schmedtmann and Campagnolo, 2015). In addition, the heterogeneity in terms of agro-climatic conditions of the area of interest and field sizes might impact the performance of the approach, as it has mostly been applied to rather homogeneous areas with large fields (Whelen and Siqueira, 2017, 2018; Xu et al., 2019).

Therefore, the main objective of this study is to investigate the performance of a supervised crop classification approach based on crop temporal signatures obtained from Sentinel-1 time series on a challenging case study with a large number of classes and a high heterogeneity in terms of agro-climatic conditions and field sizes. Hence, the temporal signatures of all crop categories were interpreted and the confusion with other classes was evaluated. The study is framed on a real case study of crop identification for the aid of CAP inspections focused on the Spanish province of Navarre, a region of strong agro-climatic and geographic diversity with pervasive cloud cover. The region was stratified into different agro-climatic regions and different classification models were trained to evaluate the benefits of stratification. Thus, a specific objective of this study was to test whether the differences in crop growing conditions by agricultural region had any influence on the crop classification accuracy. Furthermore, the relation between field size and classification results was investigated. Special attention was also paid to the analysis of results obtained with classification models based on different sets of input data, comparing VH and VV polarizations, but also, the VH/VV ratio and implementing an ensemble approach that jointly considered VH, VV and VH/VV temporal signatures.

3.2. Materials and Methods

3.2.1. Study area

The study area covers the agricultural areas of the province of Navarre (Northern Spain). This province has an extension of 10,391 km^2 and it is characterized by its landscape and climate diversity. The Northern area is mountainous, wet and forests and grasslands are predominant, while the Southern area is dryer and constituted mainly by extensive plains of the Ebro basin, where the proportion of arable land is higher with both irrigated and rainfed agriculture. Between these two areas, there is a transition zone with mixed characteristics. Due to this diversity, Navarre is stratified in seven agricultural regions with distinct agro-climatic conditions (BON, 1998) (Table 3.1).

Table 3.1. Characteristics of the agricultural regions of Navarre. *Climate classes: O (Oceanic), S (Subalpine), TOM (Transition Oceanic-Mediterranean), M (Mediterranean), CM (Continental-Mediterranean).

Region	Climate			Land cover (%)				Area (km²)	Elevation (m asl)
	Class	Annual Precip. (mm)	Average Temp. (ºC)	Rainfed crops	Irrigated crops	Forests and shrubs	Non productive		
Region 1	O*	1700	13.8	16%	0%	81%	3%	1903	7 - 1443
Region 2	S*	1500	8.7	14%	0%	84%	2%	2312	225 - 2428
Region 3	TOM*	900	12.5	47%	1%	43%	9%	778	327 - 1254
Region 4	TOM*	930	12.0	41%	2%	54%	3%	1542	324 - 1456
Region 5	M*	655	12.7	44%	7%	47%	2%	1302	304 - 1260
Region 6	M*	350	14.0	41%	30%	25%	4%	1226	259 – 652
Region 7	CM*	385	13.5	31%	33%	29%	6%	1329	226 -744

Different types of crops are grown in each region. Grasslands prevail in northern regions (1 and 2). Winter cereals are found throughout the whole province, whilst other winter crops, such as legumes and rapeseed, are common only in central regions (3 and 4). The diversity of crops in southern regions (6 and 7) is larger, where alfalfa, corn, asparagus, horticultural crops and permanent crops are also widespread. Corn, the most common summer crop, is mainly grown under irrigation systems, except in Region 1, where the high amount of precipitation allows rainfed cultivation. Other summer crops cultivated in Navarre are sunflower, found in every region, and rice, only grown in southern regions. It is important to mention that the proportion of fallow fields varies strongly depending on the region, being particularly high in Southern Navarre, especially in region 7.

3.2.2. Sentinel-1 data

The SAR data used in this study were Sentinel-1 C-band Interferometric Wide (IW) swath mode images downloaded as level-1 Ground Range Detected (GRD) products. All available Sentinel-1 images covering the study area from September 1, 2015 to December 30, 2016 were acquired (Figure 3.1). The images belonged to three relative orbits: one ascending node and two descending nodes (103ASC, 8DESC and 81DESC). Yet only the 103ASC relative orbit covered the entire area of Navarre, whereas the descending nodes excluded a small area of the province from the image, specifically the Western part in 8DESC and the Eastern part in 81DESC (Figure 3.2).

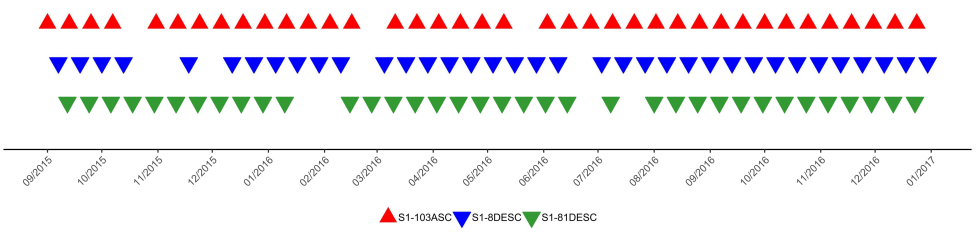

Figure 3.1. Acquisition dates of the Sentinel-1 observations used

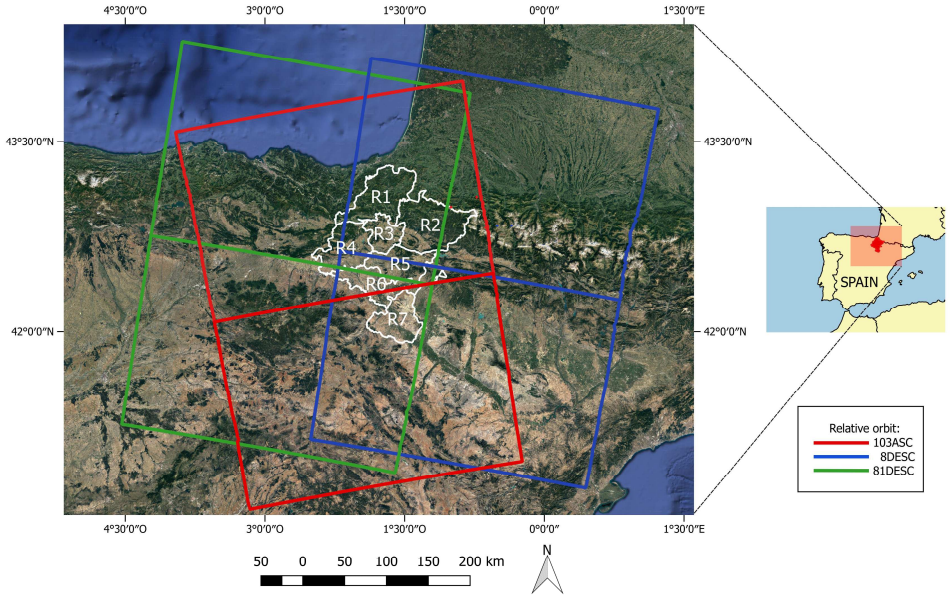

Figure 3.2. Location of the province of Navarre and its seven agricultural regions (in white, R1-R7) in Northern Spain superposed to optical images from Google. The rectangles represent the footprints of the different relative orbits of Sentinel-1 imagery covering Navarre.

The images were processed using SNAP Sentinel-1 Toolbox following these steps: (1) thermal noise removal; (2) slice assembly; (3) apply orbit file operator; (4) calibration to β^0; (5) speckle filtering; (6) terrain-flattening; (7) range-doppler terrain correction and (8) subset to the extent of Navarre. The speckle reduction was performed using a 3x3 window Gamma-MAP filter. The SRTM 1sec HGT DEM was used for terrain-flattening and terrain correction. After these corrections, terrain-flattened γ^0 backscatter coefficient values were obtained. The results of this process were 110 images with three channels each: two backscatter coefficients (VH and VV) and their

ratio (VH/VV) in dB units. The pixel size of the final products was set to 20 m. All the images corresponding to the three relative orbits were stacked according to their time.

3.2.3. Ground truth and data extraction

Ground truth information for training and validation was extracted from a geographical database containing field boundary polygons for EU Common Agricultural Policy (CAP) declarations (335,094 polygons) and inspections (21,850 polygons) for the agricultural year 2016. This database was provided by the Agriculture Department of the Government of Navarre (information not publicly available). Field inspections were carried out from June to August 2016, as part of the verifications of CAP subsidy declarations.

The geographical database was refined as follows. A 5 m inside buffer was applied to the agricultural field boundaries in order to avoid mixed pixels that could contain information from different crops. Fields smaller than 0.5 ha were masked out in the declarations dataset. In the inspections dataset, the mask was less restrictive, leaving out fields that contained less than 6 pixels. Then, field statistics were computed, obtaining the median, mean, minimum, maximum and standard deviation backscatters values for the S-1 channels. For each field, a temporal signature with the information of the three relative orbits was obtained. Fields not covered in one of the descending orbits had missing values for those specific dates. Finally, the crops were grouped into 14 classes: alfalfa, asparagus, barley, corn, fallow, grasslands, legumes, oats, other crops, permanent crops, rapeseed, rice, sunflower and wheat.

At most, 10% of farmer's declarations are expected to be wrong due to errors, fraud, etc. (Navarre, 2019). These wrongly declared fields could affect the training phase by introducing erroneous information in the learning process of supervised classification algorithms. In order to minimize the impact of wrongly declared parcels in the declarations dataset, the 10% of fields whose median backscatter temporal signature differed the most with the median signature of its declared class in each region was discarded. The median was chosen to further reduce the effect of speckle inside the field. The comparison between the temporal signature of the crop and that of each field was based on the R^2 metric. The final number of fields in Navarre is shown in Table 3.2 and Figure 3.3 (see Supplementary Materials for the number of fields per region: Table S1).

Table 3.2. Ground truth data. Number of agricultural fields and their area in the declarations and inspections datasets for the whole province of Navarre.

Crop	Declarations				Inspections			
	Nº of fields	Area (ha)	Area (%)	Mean field area (ha)	Nº of fields	Area (ha)	Area (%)	Mean field area (ha)
Alfalfa	1747	3066	1.1%	1.76	397	482	1.9%	1.21
Asparagus	430	860	0.3%	2.00	61	84	0.3%	1.38
Barley	21711	52427	19.1%	2.41	2705	5048	20.3%	1.87
Corn	3814	10693	3.9%	2.80	589	1682	6.8%	2.86
Fallow	11889	23346	8.5%	1.96	3089	4785	19.2%	1.55
Grasslands	25607	81230	29.6%	3.17	512	823	3.3%	1.61
Legumes	3776	11544	4.2%	3.06	637	1211	4.9%	1.90
Oats	2969	7336	2.7%	2.47	268	478	1.9%	1.79
Other crops	3025	10094	3.7%	3.34	838	1921	7.7%	2.29
Permanent crops	8329	13342	4.9%	1.60	2010	2190	8.8%	1.09
Rapeseed	1250	3319	1.2%	2.66	226	544	2.2%	2.41
Rice	637	1415	0.5%	2.22	72	125	0.5%	1.73
Sunflower	799	2382	0.9%	2.98	102	311	1.2%	3.05
Wheat	20722	53502	19.5%	2.58	2882	5216	20.9%	1.81
Total	106705	274556	100.0%	2.57	14388	24900	100.0%	1.73

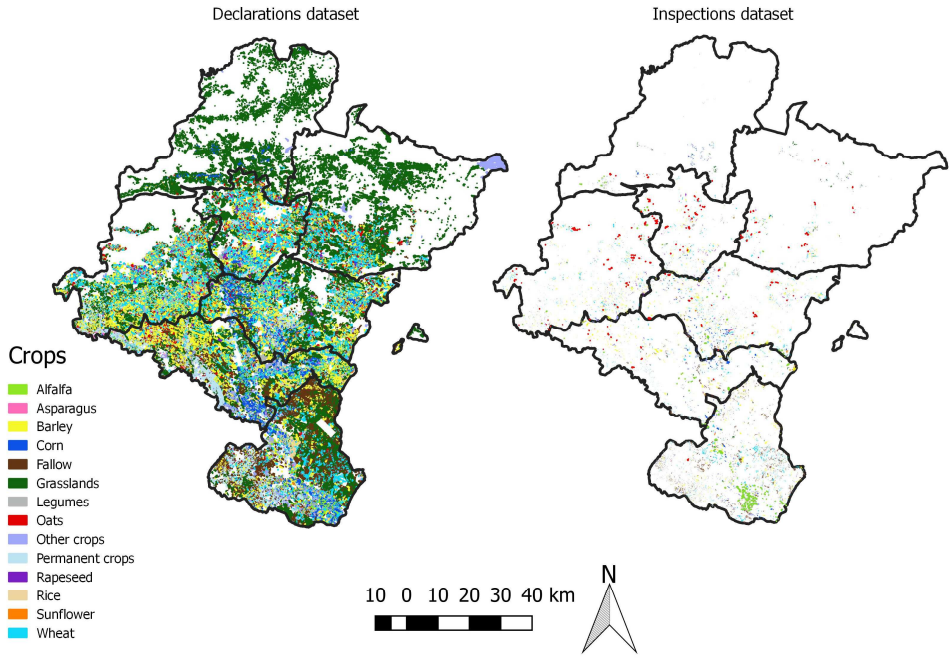

Figure 3.3. Distribution of ground truth datasets in Navarre: declarations dataset (left) with 106,705 fields, and inspections dataset (right) with 14,388 fields.

3.2.4. Separability analysis

As a preliminary step, a feature importance analysis was conducted comparing the statistical features calculated for the different polarizations. For this analysis, the Jeffries–Matusita (JM) distance was evaluated (Richards and Jia, 1999), which provides a measure of separability between pairs of classes and has been identified as a good indicator of crop separability (Van Niel et al., 2005). The JM distance between a pair of probability distributions is given by the following:

$$JM_{ij} = \left\{ \int \left[\sqrt{p\left(x|\omega_i\right)} - \sqrt{p\left(x|\omega_j\right)} \right]^2 dx \right\} \qquad (3.1)$$

where JM_{ij} is the JM distance between a pair of classes ω_i and ω_j, x is the feature observed (field backscatter statistics) and p $(x|\omega_i)$ and p $(x|\omega_j)$ are the conditional probability functions for x given ω_i and ω_j, respectively. JM distance values range between 0 and 2, with higher values indicating higher separabilities between pairs of crops.

The JM distances were calculated per date for each pair of crops, for the different polarizations and for the different statistical features. A mean JM distance value was calculated for the whole period of study, averaging the distances obtained for each date. This mean JM-distance was used to compare the importance of the different polarizations and statistical features calculated. Two statistical features obtained the highest separabilities: the mean and the median. The median was chosen for the classification because of its higher robustness to outliers.

3.2.5. Classification

A supervised classification algorithm based on the temporal signatures of crops was applied. The algorithm is similar in concept to the one used in Whelen and Siqueira (Whelen and Siqueira, 2017, 2018), and Xu et al. (Xu et al., 2019b). However, there are some significant differences in the way similarity is measured (see below). The basic concept behind this classification algorithm is that crops exhibit physical differences during the growing season due to their phenological development, which leads to different backscatter time series. Thus, the characteristic time series of each crop, as observed by Sentinel-1, is referred to as its temporal signature. This concept of temporal signature somehow mimics the idea of the spectral signature of a crop, i.e., the typical reflectance response of a crop with respect to wavelength. Therefore, we could define the temporal signature as the typical backscatter response of a crop with respect to time. The basic hypothesis is that these temporal signatures are representative of

each crop and can be used to identify whether an unknown field belongs to one class or another. In the training phase, the temporal signatures of each crop class were obtained following a three-fold cross-validation scheme using the declarations dataset, where 2/3 of the ground truth samples were used for training and the remaining 1/3 for validation, in three successive random folds. Backscatter time series were formed by adding the information obtained at the three relative orbits used. After this training phase, each field was classified by comparing its temporal signature one by one with the different crop temporal signatures. In fact, each field was assigned to the class providing the best fit between its temporal signature and the crop temporal signature. To quantify the best fit, two variants were explored, the first variant assumed the best fit as the one with the lowest root mean squared error (RMSE), and the second as the one with the highest determination coefficient (R^2). This differed from related studies, Whelen and Siqueira (Whelen and Siqueira, 2017, 2018) used a normalized error metric, and Xu et al. (Xu et al., 2019b) used a spectral similarity measure. Here, an error metric (RMSE) and a correlation metric (R^2) were evaluated, as they provide a complementary way to quantitatively compare two curves. Since the crops can have different calendars, the duration of the curves to be compared was adjusted to each crop type (Table 3.3). In addition, each crop could be characterized by three different temporal signatures depending on the channel used: VH, VV and VH/VV. Thus, three different classification schemes were applied by considering these channels individually. One additional scheme was explored by considering the similarity to the three polarizations conjointly and named as the ensemble scheme (Ens). Therefore, in total, 8 variants were explored (4 input features x 2 goodness-of-fit measures).

Table 3.3. Period of duration of the curves to compare based on crop types.

Crops	Period of comparison
Barley, legumes, oats, rapeseed and wheat	September-2015 to July-2016
Corn, rice and sunflower	April-2016 to December-2016
Rest of crops	September-2015 to December-2016

In order to test whether the differences in crop growing conditions by agricultural region had any influence on the crop classification accuracy, the supervised classification algorithm was fitted at different spatial extents: first, for the whole province of Navarre and next, stratified for each of the seven agricultural regions. In the first case, the temporal signatures were obtained for the whole province, whereas in the second, specific temporal signatures per region were obtained.

For each classification scheme, the accuracy was assessed through a confusion matrix whereby the Overall Accuracy (OA) was calculated, as well as each crop's Producer's Accuracy (PA), User's Accuracy (UA) and F1-score (the harmonic mean between PA and UA). In fact, a two-level accuracy assessment was implemented. The

first level corresponded to the 3-fold cross-validation and used the 1/3 ground truth data set left out for validation on each of the 3-folds. Additionally, a more rigorous second level validation was performed based on the field inspections dataset available, called external validation. This second dataset was obtained through field visits, so its reliability is believed to be higher. However, the sampling did not follow the standard recommendations for accuracy evaluation (Olofsson et al., 2014), and had a higher proportion of fallow, permanent crops and other crop classes (Table 2), categories that can be more difficult to classify than others such as grasslands with a much lower number of samples. This might negatively impact the OA values obtained. In addition, the size of fields was smaller compared to the declarations dataset (Table 2). Therefore, the results should be interpreted bearing this point in mind. The accuracy metrics for these two validation levels were reported in the results. Due to the small size of several fields in the external validation dataset, accuracy results were also reported and analyzed by field size considering three size groups: (1) < 0.5 ha; (2) 0.5 ha—1 ha; (3) ≥ 1 ha.

3.3. Results

3.3.1. Separability analysis

Table 3.4 shows the results for the statistical features importance analysis, where the mean JM distances for all pairs of crops was calculated. The standard deviation time series obtained the lowest values, while the mean and median achieved the highest JM distances, and were very similar. As indicated in Section 2.4., the median time series was selected for computing the temporal signatures of each crop class due to its higher robustness to outliers.

Table 3.4. Mean JM distances for all the pairs of crops for each statistical feature and polarization. The JM distances reported correspond to the average values for the whole period of study.

Polarization	Max	Min	Mean	Median	St
VH	0.20	0.23	0.29	0.29	0.08
VV	0.17	0.23	0.29	0.29	0.06
VH/VV	0.18	0.18	0.32	0.31	0.03

The separability values between each pair of crops for the median time series are represented in Figure 3.4. Although the values are, in general, low (< 1), it must be pointed out that this is an average value for the whole period of study, so they should be used for comparison between crop pairs and polarizations. There were pairs of crops that obtained similar values (in most cases low) for the three polarizations. In VH polarization, rapeseed and rice had relatively high separabilities with most of the crops. In VV, alfalfa was the crop with the highest values, followed by oats and barley. In the

case of VH/VV, the highest separabilities were achieved when comparing winter versus summer crops.

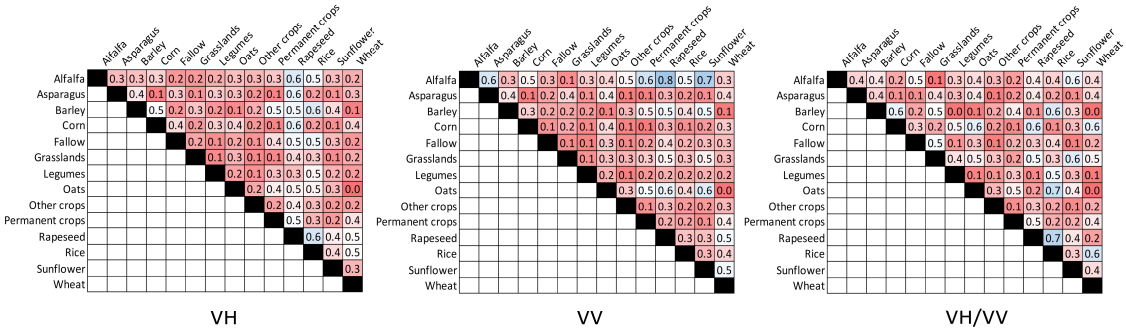

Figure 3.4. Mean JM Distance per crop pair for the median backscatter time series of the different polarizations during the whole period of study.

The temporal evolution of crop separability provides interesting information on the time periods that contribute the most to crop identification (Figure 3.5). Monthly boxplots of JM distances for all crops show that in general the period between March and August provided the highest separabilities for the three polarizations (Figure 3.5a-c). When evaluating the separability between winter crops (barley, legumes, oats, rapeseed and wheat), both VH and VV polarizations had rather high values in spring months, while VH/VV distances were considerably lower (Figure 3.5d-f). In the case of summer crops (corn, rice and sunflower) separabilities were lower than for winter crops, yet they increased between May and August for VH and VV polarizations, and between July and November for VH/VV (Figure 3.5g-i). Permanent crops had in general low separability values throughout the whole period, taking slightly higher values in winter months for VH and VV polarizations, and in spring for VH/VV (Figure 3.5j-l).

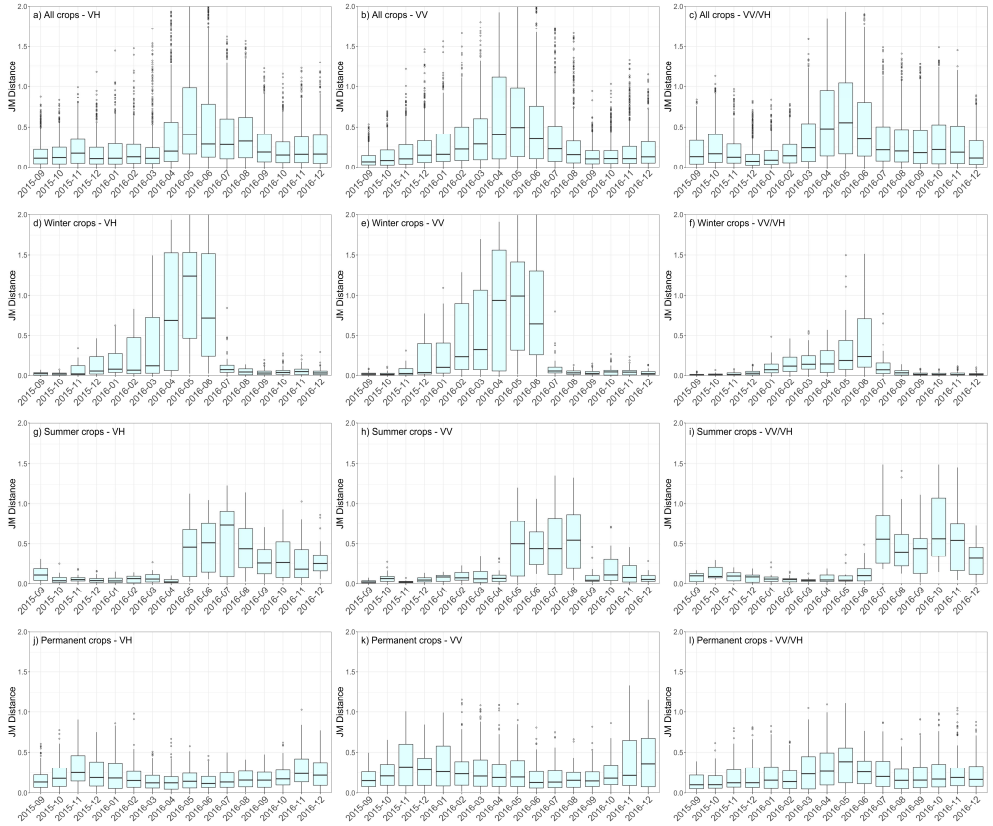

Figure 3.5. Monthly boxplots of JM distance for the three polarizations (VH, VV and VH/VV), reflecting the separability between all crops (a, b, c), winter crops (d, e, f), summer crops (g, h, i), and permanent crops (j, k, l).

3.3.2. Temporal signatures of crops

Figures 3.6–3.8 represent the temporal signatures of the studied crops for the three backscatter channels (VH, VV and VH/VV) from September 2015, 1st (day 0) to December 2016, 30th (day 487). Generally, the signature of each crop in the different regions was similar, so a general description is provided and the regions where the response differed are pointed out (see Supplementary Materials for detailed temporal signatures per region: Figure S1). For each polarization, the temporal curves of the three relative orbits had similar backscatter values and the trend was the same throughout the whole period of study. In the following sections, the temporal signatures are commented on by crop type.

a) Winter crops (barley, legumes, oats, rapeseed and wheat)

Winter cereals (barley, oats and wheat) are normally sown between October and November, with the exception of Region 7, where short cycle varieties of barley and wheat are also sown in January and February. Harvest takes place normally during June and July, starting earlier in the most southern regions. VH/VV signatures of winter cereals responded positively to the growth of vegetation along the season until the ripening stage (Figure 3.6), which occurred in May (~250 day) for barley and in June (~275 day) for wheat and oats. Then, during senescence, VH/VV values decreased until the crop was harvested. It is possible to identify in the signature of barley that harvest occurred earlier than in wheat or oats (Figure 3.6).

VH and VV temporal signatures were very similar for the three cereal crops during autumn and winter, with no clear trends and a strong variability between successive dates. In April (~200 day), a consistent decrease in backscatter was initiated in both polarizations, being sharper in VV. After ~250 days, backscatter values increased again. In this phase, barley increased more than wheat and oats. During senescence, VH values in the three cereal crops and VV in barley decreased steadily until the harvest was finalized. The temporal signatures of cereals showed the same trends in all agricultural regions, except for Region 7, where VH and VV backscatter signatures were noisier (Supplementary Materials: Figure S1).

The legume crops cultivated in Navarre are mostly beans and peas. These crops, which are intended for fodder in the majority of the region, are grown from October to July. In this case too, VH/VV responded positively to crop growth (Figure 3.6). Regarding VH, an increase in backscatter values was detected from the end of autumn due to the development of the crop. The onward increase of backscatter finished in June (~275 day), and was followed by a decrease due to crop senescence. In the case of VV, there was not a clear trend until March (~200 day), when backscatter started growing smoothly until June. VH backscatter values grew slowly in winter and early spring. There was a stronger increase from April (~215 day) to June (~275 day). After this peak, the values dropped, matching the senescence period. The temporal signatures of legumes were shorter in Region 7 (see Supplementary Materials: Figure S1). In this region, the increase of the curves came later—at the end of March—and the slope of the decreasing curves was steeper.

Rapeseed is mostly grown in the central area of Navarre. Its agricultural calendar is similar to that of winter cereals, with sowing in September and harvest in July. VH/VV signatures followed the growth of the crop along the cycle, reaching their highest peak in May (~250 day) (Figure 3.6). VH backscatter values of rapeseed increased since the

beginning of the cycle, but there was a steeper increase between the end of April (~240 day) and the peak at the end of May, coinciding with the inflorescence emergence. After this peak, VH backscatter dropped significantly until mid-July, when the crop was finally harvested. In the case of the VV signature, no clear trend was apparent until mid-April (~230 day), when the values first decayed and then suddenly increased, depicting a peak similar to the VH curve.

b) Summer crops (corn, rice and sunflower)

These three crops are usually sown in April, so the variability of the time series before this date was high in all cases (see the interquartile ranges in Figure 3.7). Harvesting of sunflower occurs around September before cereal is sown, as it is a common crop for cereal rotations. Rice and corn harvest dates can range from October to December.

Looking at VH/VV time series (Figure 3.7), the temporal signature responded positively to the growth of corn and rice, but for sunflower, it was not so evident, particularly in descending orbit curves. For this crop, backscatter values increased both in VH and VV from the beginning of June until August. In this case, VH/VV did not adequately resemble crop growth. In corn, there was an increase in VH backscatter from plant emergence until July. After that moment, there were no clear trends. Corn VV series did not show any clear trend during the growing cycle. Rice is flooded before sowing, and this caused an evident effect in VH and VV time series. There was an increase in backscatter at the beginning of April, coinciding with soil tillage, followed by a strong decrease due to the specular reflection caused by standing water in the fields. Once the crop started to grow, backscatter also increased steadily until August (~336 day). After that point, VH backscatter continued increasing until the end of the period of study, whereas VV remained steady till November (~400 day), when it started to grow. It is remarkable that the VH/VV time series of corn were all similar throughout the different regions, except for the northern regions, which were shorter than the rest (see Supplementary Materials: Figure S1).

Figure 3.6. Median temporal signatures of the main winter crops (barley, oats, wheat, legumes and rapeseed) for VH, VV and VH/VV. The three relative orbits are represented in different colors (see the legend). The interquartile range is represented for all the different time series as a colored shadow. The horizontal axes represent the period of study, from September 1, 2015 (day 0) to December 30, 2016 (day 487).

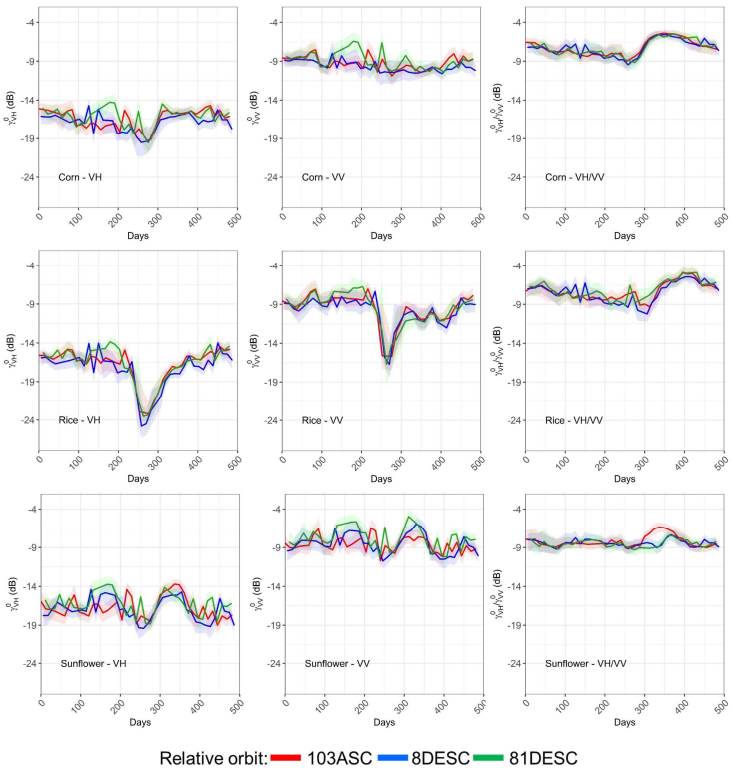

Figure 3.7. Median temporal signatures of main summer crops (corn, rice and sunflower) for VH, VV and VH/VV in Navarre. The three relative orbits are represented in different colors (see the legend). The interquartile range is represented for all the different time series as a colored shadow. The horizontal axes represent the period of study, from September 1, 2015 (day 0) to December 30, 2016 (day 487).

c) Rest of crops (alfalfa, asparagus, grasslands, permanent crops, other crops and fallow)

Alfalfa is a crop that can be sown in different periods of the year (September or spring months), whose management normally includes irrigation and several cuts throughout the cycle. These cutting events complicate the interpretation of the crop signature since different fields might be cut on different dates, resulting in a poorly informative average curve. Yet, VH and VV temporal signatures showed different peaks during the period of study, but the values remained relatively constant. A small increase in VH/VV was observed from January to May.

Asparagus follows a cycle of two years. This crop is planted between February and April of the first year, and harvested between March and May of the second year. VH and VV temporal signatures showed many backscatter peaks along the period of study. It is possible to detect an increase of VH from May to November, the period of vegetative growth (Figure 3.8). VH/VV also augmented in this period, although it increased very smoothly.

Backscatter signatures of grasslands were heterogeneous and did not follow any defined pattern (Figure 3.8). In this case also (as in alfalfa), the influence of individual cuttings complicated the interpretation of the temporal signature.

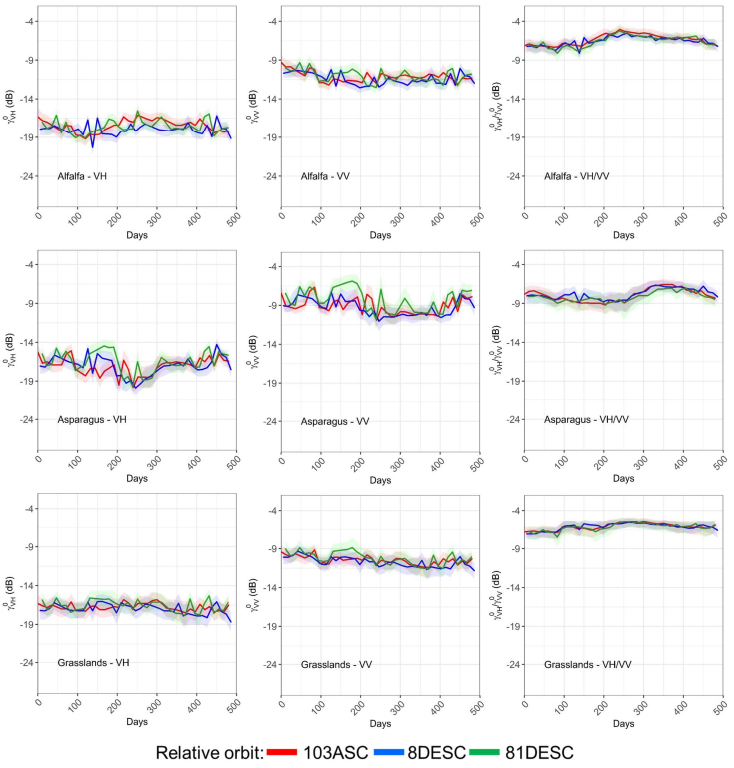

Figure 3.8. Median temporal signatures of alfalfa, asparagus and grasslands for VH, VV and VH/VV in Navarre. The three relative orbits are represented in different colors (see the legend). The interquartile range is represented for all the different time series as a colored shadow. The horizontal axes represent the period of study, from September 1, 2015 (day 0) to December 30, 2016 (day 487).

Permanent crops is another heterogeneous group including different woody crops: vineyards, olive trees, almond trees and different types of fruit trees. In VH and VV curves, there were backscatter peaks that were smoothed during spring and summer. VH/VV slightly increased from mid-spring until mid-autumn.

The great variability of crops included in the "other crops" class was clearly visible in the dispersion of the class' temporal signature (Figure 3.9). This class included herbaceous and horticultural crops with both summer and winter cycles. There were no clear trends and the interquartile range was large throughout the period of study.

Fallow temporal signatures also had high interquartile ranges. For all the three polarizations, there were no clear trends, with different peaks along the period of study.

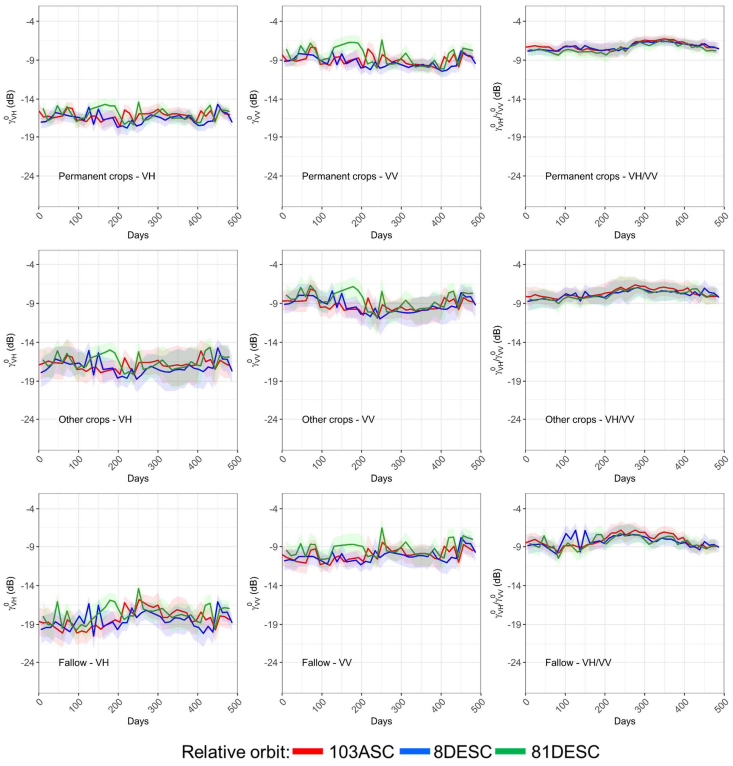

Figure 3.9. Median temporal signatures of permanent crops, other crops and fallow for VH, VV and VH/VV in Navarre. The three relative orbits are represented in different colors (see the legend). The interquartile range is represented for all the different time series as a colored shadow. The horizontal axes represent the period of study, from September 1, 2015 (day 0) to December 30, 2016 (day 487).

3.3.3. Classification results

a) General results

Overall Accuracy (OA) results showed differences among regions, polarization channels used as input, best fit metric and type of validation (cross-validation or external validation) (Figure 3.10). The best classification results were achieved in Region 3 in cross-validation, using Ens as the input and R^2 as the best fit metric, with OA higher than 85%. On the other hand, using only VH/VV as input and R^2 resulted in the poorest results, with an OA of around 50% in specific regions. In Figure 3.11, the classification map obtained with the best configuration (Ens scheme using R^2 as the best fit metric) is represented.

External validation provided systematically lower OA values than cross-validation, with the exception of Region 3, where the results were very similar. Although the accuracies were lower in the external validation, it can be observed that the results followed a similar distribution to the cross-validation output (Figure 3.10). In the following sections, the cross-validation results will be analyzed (See Supplementary Materials for the detailed external validation results).

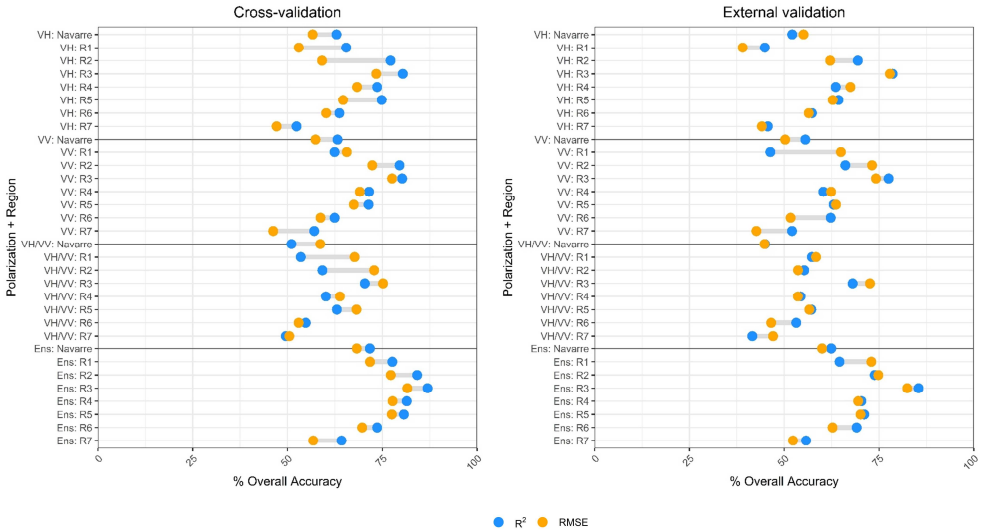

Figure 3.10. Overall Accuracy of cross-validation classification (left) and external validation classification (right).

b) Input features and goodness-of-fit metrics

OA results showed that classifications based on single polarization channels did not reach satisfactory results (Figure 3.10). VH and VV channels achieved similar accuracies in each region individually (at best 80% OA), while classifications using the VH/VV ratio as the input obtained lower values (at best, 75% OA). On the other hand, considering all the available input features (Ens in Figure 3.10) yielded the best results for all the regions, reaching 87% of OA in Region 3.

Regarding the goodness-of-fit metric used for classification, R^2 obtained the best results for VH, VV and Ens, while RMSE was the best for VH/VV. Although R^2 achieved the best global results for the most favorable schemes (Ens), there were differences among crop classes. As it is shown in Table 3.5, legumes, oats, rapeseed and rice achieved an improvement of the F1-score between 6% and 18% when RMSE was considered. On the other hand, alfalfa, fallow, grasslands and permanent crops obtained better results when R^2 was used. Asparagus, barley, corn, sunflowers, other crops and wheat did not present large differences between one metric and the other.

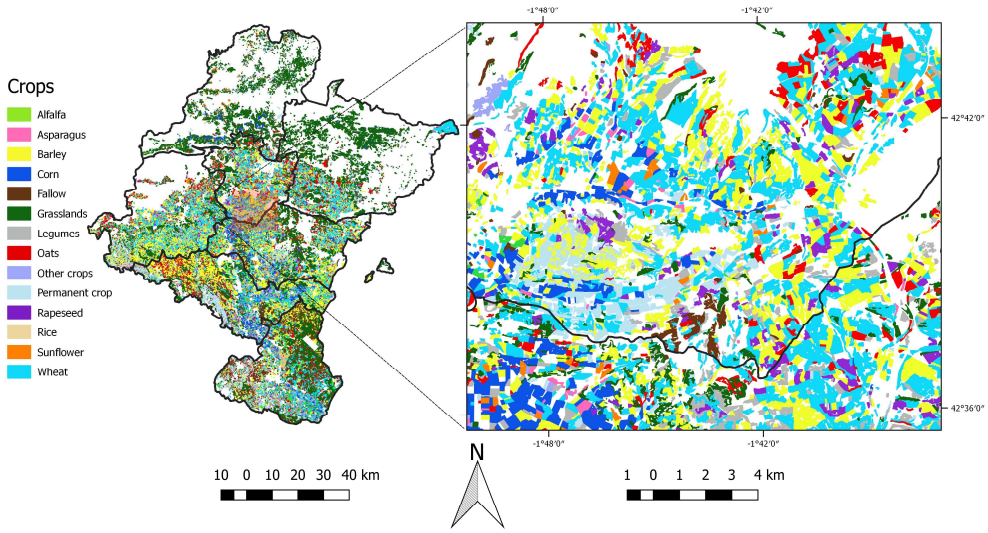

Figure 3.11. Classification map for the Ens scheme (using R^2 as best fit metric) for the different regions (left), with a detailed view of an area belonging to Regions 3 and 5 (right).

Table 3.5. F1-score of crops for Ens cross-classification scheme in Navarre using R^2 and RMSE fit metrics.

Crop class	F1-score R^2	F1-score RMSE
Alfalfa	48%	31%
Asparagus	25%	22%
Barley	87%	86%
Corn	78%	75%
Fallow	70%	62%
Grasslands	80%	71%
Legumes	47%	54%
Oats	42%	60%
Other crops	12%	14%
Permanent crops	68%	58%
Rapeseed	74%	85%
Rice	85%	91%
Sunflower	55%	53%
Wheat	76%	80%

c) Results per region

Regarding the classification results for the different regions, the OA values varied strongly. Focusing on the Ens approach with R^2 as the goodness-of-fit metric (Figure 3.10), Region 3 obtained the highest accuracies (OA = 87%). The accuracies in Region 2 were also good (i.e., 84%). Region 4 and 5 also had good OA results—around 80%. In Region 1 and Region 6, the OA results yielded 78% and 74%, respectively, whereas the worst OA results were achieved in Region 7 (64%).

Table 3.6 shows the increment in crop accuracy (PA, UA and F1-sore) when the regional results were compared to Navarre's results. The regional accuracies were extracted from a confusion matrix built by adding the individual confusion matrices of each region (Supplementary Materials: Table S2). This matrix allowed for an easy comparison of whether the stratification of the province in agricultural regions improved the general classification results. PA, UA and F1-score results were higher for the stratified case, with the exception of PA- and F1-score of fallow, which decreased by 2% and 1%, respectively, UA of grasslands decreased by 3% and PA of rapeseed decreased by 1%. In total, OA increased from 72% to 77%.

Table 3.6. Increment (Δ) in crop accuracy (PA, UA and F1-sore) when the stratified regional results are compared to Navarre's results for Ens and R^2.

Crop	Producer's Accuracy (PA)			User's Accuracy (UA)			F1-score		
	Navarre	Regions	ΔPA	Navarre	Regions	ΔUA	Navarre	Regions	ΔF1-score
Alfalfa	74%	76%	**2%**	35%	49%	**14%**	48%	60%	**12%**
Asparagus	60%	71%	**11%**	16%	25%	**9%**	25%	36%	**11%**
Barley	90%	91%	**1%**	85%	88%	**3%**	87%	90%	**2%**
Corn	75%	84%	**9%**	81%	85%	**4%**	78%	85%	**7%**
Fallow	58%	57%	**-2%**	88%	90%	**2%**	70%	70%	**-1%**
Grasslands	71%	78%	**7%**	92%	89%	**-3%**	80%	83%	**3%**
Legumes	70%	75%	**6%**	35%	48%	**13%**	47%	59%	**12%**
Oats	71%	82%	**11%**	30%	36%	**6%**	42%	50%	**8%**
Other crops	13%	19%	**5%**	12%	25%	**13%**	12%	21%	**9%**
Permanent crops	65%	72%	**7%**	72%	72%	**1%**	68%	72%	**4%**
Rapeseed	99%	97%	**-1%**	59%	70%	**11%**	74%	82%	**8%**
Rice	99%	99%	**0%**	74%	85%	**10%**	85%	91%	**6%**
Sunflower	87%	89%	**1%**	41%	47%	**6%**	55%	61%	**6%**
Wheat	69%	80%	**11%**	84%	85%	**1%**	76%	82%	**7%**

d) Results per crop

In the following section, the results per crop are provided, reporting the Ens and R^2 classification scheme for Navarre (Table 3.7) (for the rest of schemes, see the Supplementary Materials: Table S2 and Figure S2). Interesting results related to polarization channels and fit metrics are also reported.

Winter crops (barley, legumes, oats, rapeseed and wheat)

Barley is the crop that achieved the best classification results, with F1-score of 87% and balanced PA and UA results (Table 3.6). The F1-score of wheat was 71%, with higher UA than PA (84%–69%, respectively) (Table3. 6). On the other hand, oats obtained poorer classification results, with F1-score of 42%. Although PA was relatively high (70%), UA was low (30%), due to many wheat fields being incorrectly assigned to oats (Table 3.7). In general, the results obtained per region were similar (Supplementary Materials: Table S2 and Figure S2). Region 2 and Region 3 achieved the highest PA, UA and F1-score for the three cereal crops compared with other regions. In Region 1, although most barley fields were correctly classified, the F1-score was worse due to the incorrect assignment of grasslands fields to the barley class. The lowest classification accuracies of winter cereals occurred in Region 7, where other crops were also misclassified as barley and wheat, and many fallow fields were incorrectly classified as winter cereals.

Table 3.7. Confusion matrix of Ens and R^2 classification in Navarre. Values are given as a percentage of the total number of fields classified.

		Ground truth													Total (%)
		Alfalfa	Asparagus	Barley	Corn	Fallow	Grasslands	Legumes	Oats	Other crops	Perm. crops	Rapeseed	Rice	Sunflower	Wheat	
Predicted class	Alfalfa	1.21	0.00	0.05	0.05	0.35	1.09	0.04	0.00	0.11	0.45	0.00	0.00	0.01	0.06	3.42
	Asparagus	0.00	0.24	0.00	0.21	0.08	0.04	0.00	0.00	0.16	0.78	0.00	0.00	0.00	0.00	1.52
	Barley	0.01	0.01	18.25	0.07	0.76	0.37	0.15	0.18	0.24	0.01	0.00	0.00	0.01	1.45	21.50
	Corn	0.00	0.06	0.00	2.68	0.07	0.04	0.02	0.00	0.30	0.11	0.00	0.00	0.01	0.01	3.30
	Fallow	0.05	0.00	0.01	0.01	6.49	0.39	0.08	0.01	0.06	0.24	0.00	0.00	0.00	0.04	7.39
	Grasslands	0.04	0.01	0.02	0.08	0.53	17.09	0.10	0.01	0.14	0.46	0.00	0.00	0.01	0.02	18.52
	Legumes	0.22	0.00	0.74	0.05	0.94	1.24	2.47	0.06	0.30	0.14	0.02	0.00	0.02	0.79	6.98
	Oats	0.03	0.01	0.14	0.11	0.09	0.59	0.14	1.97	0.13	0.10	0.00	0.00	0.01	3.26	6.59
	Other crops	0.03	0.02	0.16	0.03	0.61	1.38	0.04	0.02	0.38	0.30	0.00	0.00	0.01	0.34	3.32
	Perm. crops	0.01	0.03	0.01	0.08	0.69	1.02	0.03	0.00	0.16	5.08	0.00	0.00	0.01	0.00	7.11
	Rapeseed	0.01	0.00	0.13	0.00	0.10	0.07	0.31	0.00	0.17	0.01	1.16	0.00	0.00	0.01	1.97
	Rice	0.00	0.00	0.00	0.03	0.03	0.02	0.01	0.00	0.09	0.00	0.00	0.59	0.00	0.01	0.80
	Sunflower	0.00	0.00	0.00	0.11	0.24	0.25	0.01	0.00	0.30	0.04	0.00	0.00	0.66	0.00	1.61
	Wheat	0.03	0.00	0.83	0.07	0.15	0.41	0.12	0.53	0.30	0.08	0.00	0.00	0.01	13.45	15.99
Total (%)		1.63	0.40	20.35	3.58	11.13	24.00	3.54	2.78	2.83	7.81	1.17	0.60	0.75	19.42	100.00

Legumes obtained an F1-score of 47%, with a higher accuracy in PA than in UA (70% and 35% respectively) (Table 3.6). Incorrectly classified legumes fields were assigned mainly to the rapeseed class (Table 3.7), and in a lower proportion, to winter cereals, grasslands and fallow. On the other hand, some grasslands, winter cereals (barley and wheat), fallow and other crops fields were wrongly assigned to legumes causing low UA values (Table 3.6 and Table 3.7). In the case of the different regions, the values of PA were also higher than UA for legumes (Supplementary Materials: Table S2 and Figure S2). The low values of UA were mainly due to incorrectly classified grasslands and in some regions (6 and 7), fallow fields. The exception was Region 3, where both PA and UA scores were high, achieving an F1-score of 89%.

Rapeseed fields reached high PA (99%), while UA was not as good (59%), leading to an F1-score of 74% (Table 3.6). The fields wrongly assigned to rapeseed belonged principally to legumes, other crops, barley and fallow classes (Table 3.7). Per region, UA values were higher than in Navarre (Supplementary Materials: Table S2 and Figure S2), reaching high values of F1-score, especially in Region 3 (97%). The exception was Region 6 (UA = 20%), due to the lower number of rapeseed fields and the incorrect assignment as barley, legumes and fallow fields.

Summer crops (corn, rice and sunflower)

The accuracy values of summer crops in Navarre were variable (Table 3.6). Rice fields were mostly correctly classified, reaching a PA of 99%, UA of 74%, and an F1-score of 84% (Table 3.6). Corn had an F1-score of 78%, with balanced PA and UA values (75% and 81%, respectively), and some confusion with asparagus and other crops fields. Wrongly assigned fields to corn were mainly the other crops class. Sunflower was the summer crop with the worst results (F1-score of 55%, PA of 87% and UA of 41%), with confusion with other crops and fallow fields, causing these low UA values (Table 7).

Rice, which is only present in Regions 6 and 7, achieved slightly higher accuracies when the classification per regions was considered, with F1-scores of 90% and 93% for Regions 6 and 7, respectively (Supplementary Materials: Figure S2). In the case of corn, the classification was highly accurate in Regions 3 and 5 (F1-scores of 96% and 93%). In Regions 6 and 7, where the majority of corn fields were found, the F1-scores were also good (around 83%). Yet, in Regions 1 and 2, the UA was much lower, mostly due to confusion with grasslands. In Region 4, UA values were not high (43%) due to the small number of corn fields. Sunflower repeated the patterns of Navarre in the different regions, with lower UA values compared to PA values, except in Region 3, where again, both PA and UA were high (100% and 91%) (Supplementary Materials: Table S2 and Figure S2).

Rest of crops (alfalfa, asparagus, grasslands, permanent crops, other crops and fallow)

The classification of alfalfa fields achieved an F1-score of 48%, with higher PA (74%) than UA (35%) (Table 3.6). PA's errors were mostly due to being wrongly assigned to legumes, while UA values were impacted by confusion with grasslands, permanent crops and fallow (Table 3.7). In Regions 6 and 7, where alfalfa is most frequent, the results were similar (Supplementary Materials: Table S2 and Figure S2).

Asparagus was one of the crops with the poorest classification results, with an F1-score = 25%, PA = 60% and UA = 16%. Many fields belonging to the class permanent

crops were classified as asparagus and, in a smaller proportion, corn and other crops (Table 3.7). These incorrect assignments were even higher than the actual number of asparagus fields, which explained the low UA. Similar results were obtained per region (Supplementary Materials: Table S2 and Figure S2).

Grasslands was the largest class in the declarations dataset. The classification had higher UA values (92%) than PA (71%), achieving an F1-score of 80% (Table 3.6). Although PA was not low, wrongly classified grassland fields impacted the UA of many different crops due to their higher proportion (Table 3.7). The main errors in PA were due to the confusion with other crops, legumes, alfalfa and permanent crops. Per region, similar results were obtained, except for Regions 6 and 7, where results were worse (F1-scores below 70%) (Supplementary Materials: Table S2 and Figure S2).

The class permanent crops obtained intermediate accuracies (PA = 65%, UA = 72% and F1-score = 68%). The main error in UA was related to incorrectly classified grasslands and fallow fields, while PA was affected by the confusion with asparagus class. Similar results were achieved per region (Supplementary Materials: Table S2 and Figure S2), with the exception being Region 3, where UA dropped to 29%.

Other crops was the class that was worst classified, with F1-score = 12%, PA = 13% and UA = 12% (Table 3.6). In all the regions, the results were also poor, and confusion occurred with many different classes, i.e., grasslands in northern regions, and permanent crops and fallow in southern regions.

Finally, fallow classification achieved better UA than PA (88% and 58%, respectively) and had an F1-score of 70%. The wrongly classified fallow fields were assigned to many different crop classes, but particularly to legumes, barley, other crops and grasslands. In contrast, in Region 1, the classification results were poor (F1-score = 2%), while in the rest of the regions, results were similar to those of the whole province.

e) Influence of field size

In most of the classification schemes and regions, the OA values obtained were higher for large fields in comparison with small ones. In some cases, the difference between small (< 0.5 ha) and large fields (> 1 ha) was higher than 14% (Table 3.8).

Table 3.8. Overall Accuracy for the stratified external validation dataset by field area of Ens and R2 classification.

	Field size (ha)		
	<0.5	**0.5-1**	**>1**
Navarre	51%	58%	65%
Region 1	56%	66%	64%
Region 2	57%	70%	75%
Region 3	75%	83%	87%
Region 4	60%	65%	74%
Region 5	60%	69%	72%
Region 6	62%	71%	68%
Region 7	45%	51%	60%

Regarding the results per crop, it can also be observed that, in most cases, F1-score values were higher for large- and medium-sized fields, in contrast to smaller fields (Figure 3.12). This was particularly true for fallow and rice classes, where the accuracies were considerably higher for large fields than for small ones.

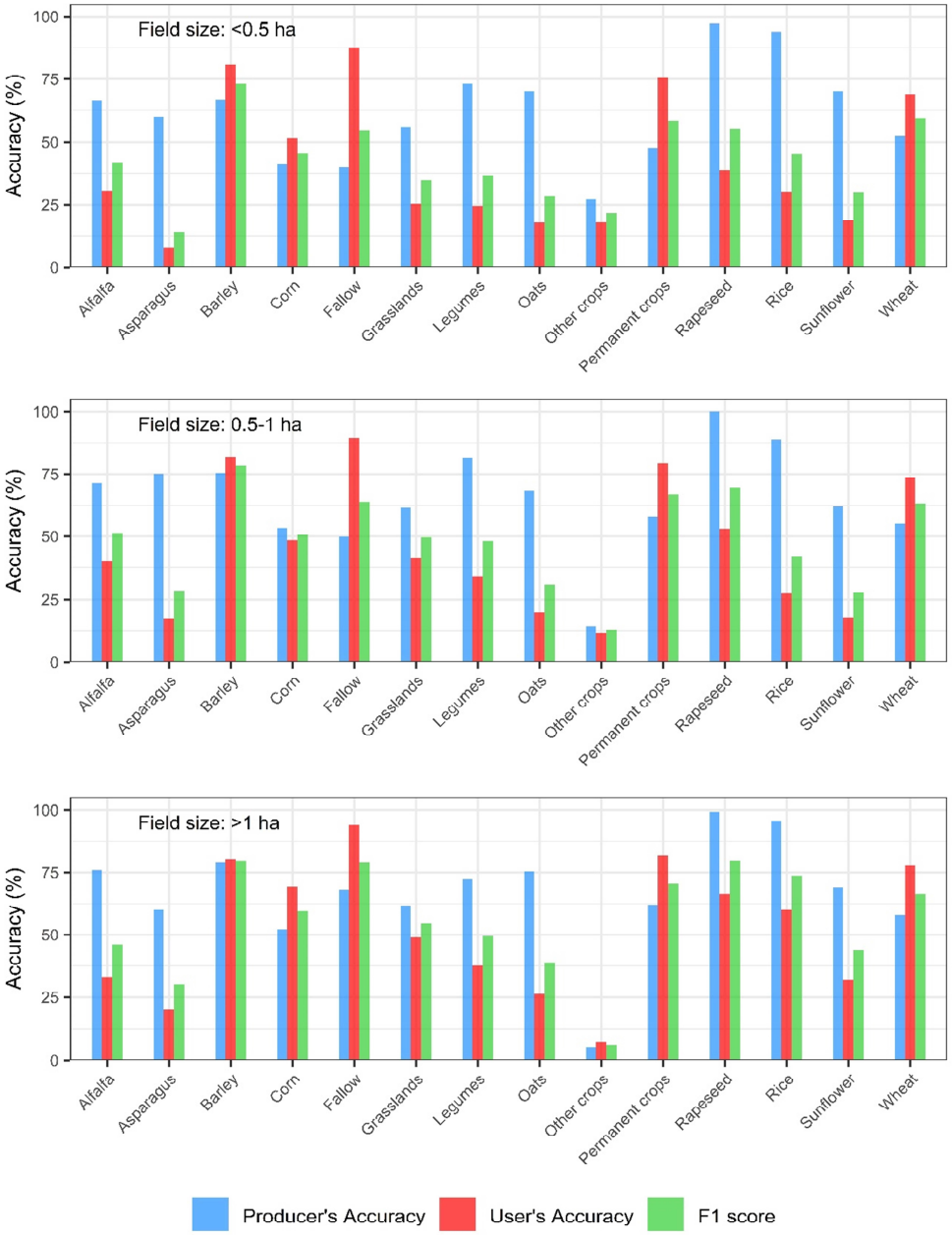

Figure 3.12. Classification map for the Ens scheme (using R2 as best fit metric) for the different regions (left), with a detailed view of an area belonging to Regions 3 and 5 (right).

3.4. Discussion

3.4.1. Crop classification results with regard to its temporal signatures

This study implemented a supervised classification technique based on the temporal signature of crops to a CAP inspection case study on a region with strong agro-climatic diversity. The results obtained varied significantly from crop to crop, with the crops that exhibited any singularities in their time series being the best classified. Cereals, and in particular, wheat and barley, are the most common crops in the region, and showed quite a distinctive temporal signature with a strong decay in backscatter in the stem elongation phase, where surface scattering from the soil was progressively attenuated by growing cereal stems (Brown et al., 2003). This decay was particularly strong in VV, due to the vertical arrangement of stems through the so-called differential attenuation mechanism (Larranaga et al., 2013; Mattia et al., 2003). This characteristic decay was followed by an increase in backscatter due to the structural changes experienced by the canopy during the heading stage (Loosvelt et al., 2012; Mattia et al., 2003; Skriver et al., 1999; Veloso et al., 2017). At this moment, a strong and well-defined peak occurred in VH and VV barley time series, probably due to the bending of barley spikes at this phase (Larranaga et al., 2013; Skriver et al., 1999), which did not occur in wheat and oats that maintained a vertical geometry. From ripening to senescence and harvest, cereal canopies dried-out and a greater penetration (Liu et al., 2013) and a higher influence of the soil surface (Mattia et al., 2003) was observed. The unique backscatter pattern of barley allowed a successful identification with F1-scores of ~85%. Wheat was also adequately identified (F1-score of ~76%), although was sometimes confused with oats, which was the cereal crop with the poorest results (F1-score of ~50%).

The architecture of legumes and rapeseed canopies is completely different from cereals, with heterogeneous shrub-like structures that produce volume backscatter (Skriver et al., 1999). This is the reason why VH and VV time series were so different from cereals, although the period of cultivation was the same. Unlike cereals, rapeseed and legumes' curves in VH started increasing right after the germination of the crop, due to the volume scattering produced by plants (Cable et al., 2014; Fieuzal et al., 2013; Larranaga et al., 2013; Yang et al., 2014). This behavior changed suddenly for rapeseed at the period of pod formation after flowering, producing a strong peak in VH (and to a lesser extent in VV). In VV, just before this peak, a small backscatter decrease was observed coinciding with the start of flowering, as also reported by Wiseman et al. (Wiseman et al., 2014) and Veloso et al. (Veloso et al., 2017). Ripening and senescence produced a strong decrease in VH backscatter (and to a lesser extent in VV) due to canopy drying and a reduction of the volume scattering component. This decrease

occurred for both rapeseed and legumes, although it was much greater for the former. Accordingly, classification results were successful for rapeseed (F1-score of ~75%) but worse for legumes (F1-score of ~50%), whose temporal signature was confounded with other crop classes.

Summer crops also achieved good classification results. Corn, with an F1-score of ~75%, was characterized by a rather insensitive VV curve, and on the contrary, a clear increase in VH due to volume scattering during the vegetative growth phase (250–300 days) until the plants reached their maximum height (Bériaux et al., 2015), after which, it remained rather constant. Rice achieved an F1-score of ~85%, mainly due to its characteristic time signature due to flooding at the time of sowing, which lead to very low backscatter due to specular reflection (Hoang et al., 2016). After rice emergence, VV increased rapidly due to double-bounce scattering, followed by a subsequent increase in VH polarization caused by volume scattering from the rice canopy that went on until the end of the season. Sunflower was the summer crop with the poorest F1-score (~60%), but this was not expected due to its rather unique temporal signatures exhibiting both VH and VV sensitivity to crop growth in the stem elongation phase where most vegetative growth occurred (250–400 days). The structure of sunflower, with broad leaves and thick stems with large open spaces between plants, caused both volume and double-bounce backscatter (Macelloni et al., 2001a). Sunflower fields were mostly correctly classified (PA > 85%), but a significant number of grasslands, fallow and other crops fields were incorrectly assigned to sunflower, leading to low UA values.

Asparagus, although having a biannual cycle, had a backscatter behavior similar to summer crops, with an increase in VH in summer (250–400 days) coinciding with the period of vegetative growth. Bargiel et al. (Bargiel et al., 2010) observed an increase in VV backscatter at X-band for this crop during summer, but our data at C-band showed a rather constant VV response. Classification results for asparagus were poor (F1-score of ~25%), mainly due to the incorrect assignment of corn and grasslands fields to asparagus (UA ~20%).

Some categories have a large variability in terms of management or species compositions. For instance, grasslands, alfalfa and fallow are classes where agricultural practices can vary significantly. Grasslands can range from an intensified management (in terms of soil preparation, sowing, fertilizing and mowing) to extensive rangelands that behave more like permanent covers and are only grazed in summer season. Fallow fields can also vary significantly depending on the management applied to spontaneous vegetation, i.e., it might be left to grow or it might be removed either chemically or mechanically. Alfalfa is a forage crop that is mowed several times throughout the season. However, the exact mowing dates might vary significantly from field to field, so the mean temporal signature might not be informative for this crop, showing a high

variability during the season (the same applies for grasslands and fallow). Surprisingly, classification results were not as poor as expected given the rather unspecific temporal signatures of these crops. The three categories achieved quite good PA values (~70% for alfalfa and grasslands and ~60% for fallow), and UA values were high for grasslands and fallow but low for alfalfa. In fact, a significant amount of grasslands and fallow fields were incorrectly assigned to alfalfa causing its UA to fall. Altogether, such high variability categories should be better addressed by designing field-specific strategies (not based on median class signatures), such as those based on the multitemporal interferometric coherence (Tamm et al., 2016).

Permanent crops was another heterogeneous class including vineyards, olives and other fruit trees, yet its classification was not too bad (F1-score ~70%). In this case, soil backscatter was predominant due to the open spaces between trees. Thus, during the period of study, the backscatter values presented peaks without a clear increasing or decreasing pattern, probably responding to soil moisture dynamics. Li et al. (Li et al., 2019) also found a stable VH signal throughout the year in almond and walnut trees. During spring and summer, backscatter variations in VH and VV were smoothed, probably due to slower soil moisture dynamics at this time of the year, and also, because the growth of leaves during this period attenuated soil backscatter.

However, the worst results were obtained by the other crops class, a very heterogeneous category composed of a range of different minor crops, with both summer and winter cycles. Again, in this case, temporal signatures did not show any clear pattern, and in turn, its variability was very large, particularly in summer months. As a result, F1-scores dropped down to ~10%, and other crops fields were confounded both with summer (i.e., corn and sunflower) and winter crops (i.e., wheat and legumes). A redefinition of this class (e.g., splitting it into two more specific categories), could lead to an improved classification of this class but also, to higher UA values for other classes that were mistaken with it.

3.4.2. General results

Previous research recommended multi-temporal optical data as the primary source for crop classification (McNairn et al., 2009a), reporting overall accuracies ~12% higher when optical data were used instead of SAR data. Indeed, crop classification approaches based on optical data provided typically accuracies above 75%, as long as imagery acquired in key phenological stages were available (Belgiu and Csillik, 2018; Defourny et al., 2019; Immitzer et al., 2016; Schmedtmann et al., 2015). Yet, optical and SAR data provide complementary information that when combined might result in enhanced classification results (Orynbaikyzy et al., 2019). However, when cloud conditions limit the viability of optical data, SAR observations still provide useful

information (Van Tricht et al., 2018; Whelen and Siqueira, 2017). The classification results obtained here were considered successful given the number of crop classes considered, the large number of fields classified and the agro-climatic diversity of the territory.

In this study, the configuration using the three S-1 features as the input (Ens) obtained the best results. For most crops, VH/VV responded positively to crop growth from emergence to senescence. Previous studies already identified VH/VV as a notable indicator of crop development correlated with fresh biomass (Veloso et al., 2017) and plant height (Canisius et al., 2018). Yet, this information was very similar for crops of the same season and did not enable their separation. Therefore, VH/VV might differentiate when a field is vegetated or not, and even the start and end of the growing season, but crops with different morphological structures might produce very similar VH/VV temporal signatures. On the contrary, VH and VV polarizations provided diverse time series depending on different backscattering mechanisms and phenological events of the crops, and thus were more useful for crop classification. In each phenological stage, the structural characteristics of vegetation elements (size, shape, density, orientation) can vary (Li et al., 2019), but also, their water content and dielectric properties (Liu et al., 2013a), in particular, in the senescence phase (Veloso et al., 2017; Whelen and Siqueira, 2017). VH and VV were useful to identify cereal heading (Moran et al., 2012) and rapeseed pod filling. Regarding backscattering mechanisms, VH was sensitive to volume scattering (e.g., in legumes, rapeseed, corn and sunflower) (Li et al., 2019; Whelen and Siqueira, 2017), and can be used as a vegetative growth indicator for these crops (Baronti et al., 1993). On the other hand, VV polarization was sensitive to surface scattering and its attenuation by cereals in the stem elongation phase (Mattia et al., 2003; McNairn et al., 2009b), and also, to double-bounce effects on sunflower (Macelloni et al., 2001) and rice (Nguyen and Wagner, 2017). Rice flooding was also easily detected in both VH and VV (Nguyen and Wagner, 2017). During the rest of the year, the main drivers of backscatter dynamics were meteorological events, soil preparation practices and harvesting procedures (Whelen and Siqueira, 2017).

3.4.3. Influence of field size

Classification results were clearly influenced by the size of the fields, with fields >1 ha showing an OA improvement of ~14% when compared to small fields (<0.5 ha). Similar results were already reported in the recent literature (Defourny et al., 2019), showing that the operational applicability of crop-type mapping based on imagery with a spatial resolution of ~20 m in regions with very fragmented landscapes requires further advances to address the small field issue.

3.4.4. Regional stratification

In general, our results improved when the classification algorithm was trained and applied separately to the different agricultural regions, with increments of ~5% in the OA and stronger improvements (10%) in the F1-score for particular crops such as alfalfa, legumes and asparagus. In our case, small and homogenous regions with a rather reduced crop legend and a small proportion of "difficult" categories (other crops, fallow, etc.) obtained the best results (i.e., Region 3). In turn, regions with a larger crop diversity, a higher proportion of fallow fields and more variability in crop management techniques obtained poorer results. In particular, southern regions (Region 6 and Region 7), due to their dryer climate, had a mix of rainfed and irrigated agriculture leading to a larger crop variety (including a higher proportion of other crops and fallow fields), more variable temporal signatures and poorer classification results. This occurred for categories such as winter cereals or legumes, where both long and short cycle varieties coexisted, some of them were intended for forage and some for human consumption. Corn was another category that improved after regional stratification, due to its diversity, with a shorter cycle rainfed corn for fodder production in Region 1 and a longer cycle irrigated corn for grain production in the rest of the province. Altogether, the results of this study showed that in geographic areas with great agro-climatic diversity, crop temporal signatures might change significantly; therefore, operational classification approaches should be applied to stratified regions with less variability.

3.4.5. Methodological aspects

Some other methodological details were found to have an influence on the results obtained. First, the use of either R^2 or RMSE as the metric for finding the best matching category did not provide consistent results in all cases. In general, the best classification strategies were those using R^2. However, we found that RMSE provided the best results when VH/VV was the only input feature used, whereas R^2 was better for VH, VV and Ens. The different shape of their temporal signatures (with smoother dynamics for VH/VV) could be behind these differences. Analyzing the results per crop, it was found that oats, rapeseed, legumes and rice obtained significantly better results with RMSE (6%–18% improvement in F1-score), when the Ens configuration was used. In contrast, alfalfa, permanent crops, grasslands and fallow had higher F1-scores with R^2, and for some cases, differences were minor (e.g., asparagus, barley, corn, sunflower or wheat). R^2 might be more insensitive to class variability, with fields with slightly delayed phenology still having high R^2 values, as long as they show a similar trend. Furthermore, RMSE might be more sensitive to the actual backscatter values and to abrupt changes in the time series due to particular phenological events or soil moisture variations.

3.4.6. Future research

In this study, the duration of the time series was adjusted to the seasonality of the crops to be compared (winter, summer and permanent crops). This aspect was crucial due to the negative influence of the periods without vegetation cover on the variability of the time series. During autumn and winter months, differences in soil preparation dates and techniques (particularly important in furrowed crops such as asparagus), and soil moisture dynamics responding to wetting and drying events (Veloso et al., 2017) are the main backscatter drivers because vegetation is still short. Furthermore, in the case of summer crops, during this period, spontaneous vegetation might show false dynamics that confound crop recognition. After harvest, some fields might exhibit regrowth of natural vegetation, while some others might be quickly tilled and prepared for the next rotation, so this period also introduces uncertainty into the classification. Future improvements of this classification approach might try to further adjust the duration of the growth season, attending to the periods were largest differences were observed between crop classes. Also, more attention should be paid in the future to the establishment of the crop legend to classify, since mixed classes such as legumes, other crops or grasslands have shown to be too heterogeneous, resulting not only in their poor identification but also, in the confusion with other categories, affecting the overall results of the classification. Finally, further research is required to compare the results obtained using radar backscatter data versus optical reflectance on crop classification strategies like this based on image time series data. Furthermore, a comparison with state-of-the-art machine learning classification algorithms (e.g., Random Forests or Support Vector Machines) will provide interesting information for the selection of the best classification approach in each case study.

3.5. Conclusions

This paper presented a supervised crop classification technique based on the temporal signatures extracted from Sentinel-1 (VH, VV and VH/VV) time series, and applied it to a large dataset, framed in the CAP inspection process in a region with high agro-climatic diversity.

The results showed that crops whose temporal signatures depicted singularities along the growing season achieved accurate classification results (F1-score ~75%), e.g., winter cereals or rice. VH and VV temporal signatures proved to be sensitive to various phenological events where the structural characteristics or the water content of the canopy varied. In contrast, VH/VV were sensitive to vegetation growth, allowing the determination of whether a field was vegetated or not but providing poor information for crop identification. The combination of VH, VV and VH/VV time series as input features provided accurate results (OA > 70%).

A detailed analysis of the results indicated that field size strongly influenced the results with large fields (> 1 ha) achieving ~14% higher accuracy than small ones (< 0.5 ha) for the same class. Agro-climatic diversity was also crucial, with results improving when classifications were stratified for local agricultural regions, in particular, for legumes, alfalfa or asparagus. Also, crop diversity, variability in terms of management techniques and a high proportion of fallow fields negatively affected the obtained results per region. The definition of the crop legend should avoid heterogeneous crop classes (e.g., other crops or legumes) that were found to be difficult to identify, and affected the UA of some classes with good PA. Whenever possible, these classes should be divided into more homogeneous ones.

Altogether, the obtained results suggest that similar approaches based on Sentinel-1 time series could be implemented operationally in regions with frequent cloud cover, in the framework of CAP inspection, or with any other purposes like crop acreage estimation or ensuring food security. Further studies comparing these results with those obtained using optical image time series, or state-of-the-art machine learning classification algorithms might confirm this conclusion.

CHAPTER 4

ON THE INFLUENCE OF ACQUISITION GEOMETRY IN BACKSCATTER TIME SERIES OVER WHEAT

Published in: Arias, M., Campo-Bescós, M.A. and Álvarez-Mozos J. 2022. On the influence of acquisition geometry in backscatter time series over wheat. International Journal of Applied Earth Observation and Geoinformation, 106, 102671. DOI: 10.1016/j.jag.2021.102671

Abstract

Dense time series of Sentinel-1 imagery are an invaluable information source for agricultural applications. Multiple orbits can observe a specific area and their combination could improve the temporal resolution of the time series. However, the orbits have different acquisition geometries regarding incidence and azimuth angles that need to be considered. Furthermore, crops are dynamic canopies and the influence of incidence and azimuth angles might change during the agricultural season due to different phenological stages. The main objective of this letter is to evaluate the influence of different acquisition geometries in Sentinel-1 backscatter time series over wheat canopies, and to propose a strategy for their correction. A large dataset of wheat parcels (~40,000) was used and 344 Sentinel-1 images from three relative orbits were processed during two agricultural seasons. The first analysis was a monthly evaluation of the influence of incidence angle on backscatter (σ^0) and terrain flattened backscatter (γ^0). It showed that terrain flattening significantly reduced the backscatter dependence on incidence angle, being negligible in VH polarization but not completely in VV polarization. Incidence angle influence in VV backscatter changed in time due to wheat growth dynamics. To further reduce it, an incidence angle normalization technique followed by an azimuthal anisotropy correction were applied. In conclusion, γ^0 enabled a reasonable combination of different relative orbits, that may be sufficient for many applications. However, for detailed analyses, the correction techniques might be implemented to further reduce orbit differences, especially in bare soil periods or winter months.

Keywords: azimuthal anisotropy, incidence angle, normalization, SAR, Sentinel-1, wheat

4.1. Introduction

Synthetic aperture radar (SAR) imagery has proven to be useful for agricultural applications (Liu et al., 2019), such as crop classification, yield forecasting or soil moisture estimation (Steele-Dunne et al., 2017). The launch of the Sentinel-1 mission in 2014 made freely available an unprecedented collection of worldwide systematically collected C-band observations (Berger et al., 2012). Currently, the mission consists of two twin satellites that allow a nominal temporal resolution of 6 days. Yet, there can be multiple orbits with different acquisition geometries (incidence and azimuth angles) observing a specific area, whose combination can reduce the revisit time to less than two days in many parts of Europe (Weiß et al., 2021). These denser time series could be interesting for several different applications. However, the combination of time series acquired with different orbits might not be so straightforward, due to the sensitivity of backscatter to image geometry, mainly incidence and azimuth angle variations (Bartalis et al., 2006; Gauthier et al., 1998; Rizzoli and Bräutigam, 2014; Ulaby et al., 1982).

In effect, observed backscatter values vary depending on the incidence angle (Ulaby et al., 1982). However, the magnitude of these variations depend on target characteristics and scattering mechanisms (Ardila et al., 2010). Smooth targets dominated by the specular component of surface scattering were found to be particularly sensitive to incidence angle variations (Skriver et al., 1999b). Conversely, very rough soils or vegetation covers, where volume scattering predominates, have a lower incidence angle dependence. Similarly, azimuthal effects are more prominent in surface scattering situations with predefined directional structures. Certainly, both incidence and azimuth angle effects increase in areas with moderate to strong topography.

Different methods exist to normalize the incidence angle influence on backscatter. One of the most common techniques is the cosine correction (Ulaby et al., 1982), based on Lambert's law for optics, that was also modified to account for the dynamics of maize (Feng et al., 2021). Other methods are based on regression analysis (Phung et al., 2020; Wagner et al., 1999), statistical techniques (Mladenova et al., 2013; Ye et al., 2015), on the backscatter and incidence angle product (Kaplan et al., 2021), on radiative-transfer models (Ardila et al., 2010) or on empirical relationships with NDVI (Fieuzal et al., 2013). Regarding the correction of azimuthal effects, Schaufer et al. (2018) found that azimuthal anisotropy was mainly caused by the orientation of topographic slopes and proposed a matching method for its correction. However, some studies identified a different azimuthal behavior for different land covers (Bartalis et al. 2006), and eventually a reduced influence of azimuthal angles for wheat when compared to

incidence angle variations (Weiß et al., 2021). On the other hand, Small (2011) proposed a radiometric terrain correction method that 'flattens' backscatter values and potentially enables the combined use of multi-track and multi-sensor backscatter time series.

The case of agricultural land-covers is peculiar, since most crops are dynamic targets, and thus scattering mechanisms change during the agricultural season. This might change the influence of incidence and azimuth angle in backscatter. Wheat is one of the main crops cultivated worldwide, with more than 200 million ha cultivated per year (FAOSTAT, 2021), and it has been extensively investigated using remote sensing data, in particular SAR data (Liu et al., 2019). Despite its global importance, the effect of incidence and azimuth angle variations in backscatter time series, and the dynamics of these effects during the growing season of wheat have not been sufficiently studied. Therefore, the objective of this letter is to evaluate the influence of incidence and azimuth angles on Sentinel-1 backscatter time series acquired over wheat fields with different orbits, and to propose a strategy for their correction.

4.2. Study area

4.2.1. Study area

The study area corresponds to the agricultural areas of the province of Navarre (Northern Spain) (N42º40'4.8" and W1º38'52.8"). This province is relatively small (10,391 km^2), but the diversity of landscape and climate conditions creates regions with marked differences in terms of cropping patterns and agricultural management strategies, and as a result the province can be divided into seven agricultural regions (Arias et al., 2020).

4.2.2. Sentinel-1 images

All available Sentinel-1A and B ground range detected (GRD) images covering Navarre from 1 September 2016 to 31 August 2018 were used for the analysis. These corresponded to one ascending node (103ASC) and two descending nodes (8DESC and 81DESC). In total, 344 images were used. The median incidence angle for 103ASC was 41^0, for 8DESC was 43^0 and for 81DESC was 34^0.

Scenes were processed using an automated pipeline in SNAP Graph Processing Toolbox that followed this process: 1) thermal noise removal; 2) slice assembly; 3) apply orbit file; 4) calibration; 5) speckle filtering (3x3 Gamma-MAP); 6) range-doppler terrain correction and 7) subset to the extent of Navarre. This process produced σ^0 backscatter values in dB units. A second processing chain was implemented including

the terrain flattening algorithm (Small, 2011) and resulting in γ^0 backscatter coefficients (in dB units too). For the terrain flattening and terrain correction steps the SRTM 1sec HGT DEM was used. The resulting images had an output pixel size of 20 m. As an additional output, the local incidence angle map was generated for each scene, and used for subsequent analyses. For conciseness, the local incidence angle will be referred to as incidence angle in the article.

4.2.3. Wheat parcels dataset

Wheat is cultivated as a winter crop in Navarre. It is typically sown in the months of October or November and harvested at the end of June or in the first week of July. All wheat parcels for agricultural years 2017 and 2018 were extracted from the EU Common Agricultural Policy (CAP) declarations database provided, as an anonymized version, by the Agricultural Department of the Government of Navarre. A 5 m inner buffer was applied to the vector file and parcels smaller than 0.5 ha were discarded. The median backscatter time series per parcel were calculated for each orbit. With the aim to exclude parcels that might be wrongly declared as wheat, 10% of parcels most dissimilar to the typical (median) wheat time series were masked out (Arias et al., 2020). The final number of wheat parcels used was 18750 for 2017 and 20374 for 2018.

4.2.4. Incidence angle influence on backscatter

As a preliminary analysis, the dependency of σ^0 and γ^0 on the incidence angle (θ_{loc}) was evaluated by means of the slope of the linear regressions $\sigma^0=f(\theta_{loc})$ and $\gamma^0=f(\theta_{loc})$ fitted for each month. For this, the two descending orbits (8DESC and 81DESC) were taken into account, covering the complete growth cycle of wheat in Navarre. There, the season starts in September and ends in August the year after. Therefore, this monthly evaluation comprised both periods of bare and vegetated soils.

Results (Figure 4.1) showed clear differences between the slope values obtained for σ^0 and γ^0 backscatter, with typical slope values in VV polarization of –0.25 dB/^{0} for σ^0 that decreased to –0.10 dB/^{0} for γ^0. For VH polarization slope values (in absolute terms) were smaller but differences between σ^0 and γ^0 backscatter were similar, with slope values of -0.18 dB/^{0} for σ^0 and -0.03 dB/^{0} for γ^0. Therefore, the influence of the incidence angle in backscatter significantly decreased when the terrain flattening process was applied. Yet, some residual influence seemed to be present.

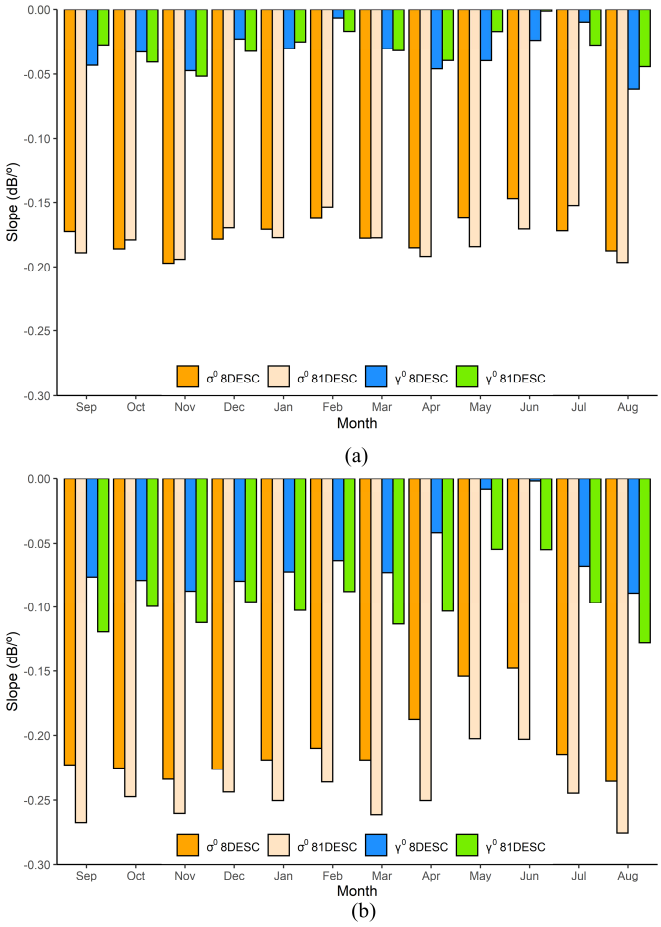

Figure 4.1. Monthly slope for wheat parcels. The slope accounts for the linear relationship between backscatter and incidence angle. (a) Slope for VH polarization. (b) Slope for VV polarization.

There were clear differences between VH and VV polarizations (Fig. 1). The slope values were much smaller for VH. The mean slope value of -0.03 dB/^{0} for γ^0, indicated an almost negligible influence of the incidence angle in γ^0 for this polarization. In VV, the obtained slope values were larger, illustrating that the influence of the incidence angle on backscatter remained after terrain flattening. Slope values were slightly larger for the 81DESC orbit that had lower incidence angles. For VV polarization, slope values varied during the year. The largest values were achieved in periods of smooth bare soils (e.g., November after sowing or August after harvest) (~-0.23 dB/^{0} for σ^0 and ~-0.10 dB/^{0} for γ^0) that decreased steadily with wheat growth reaching a minimum in May and June (~-0.16 dB/^{0} for σ^0 and ~-0.04 dB/^{0} for γ^0). In these months, wheat canopy was at

its maximum (BBCH growth stages 5-8) and the incidence angle influence was negligible.

This preliminary analysis recommends applying the terrain flattening algorithm whenever incidence angle variations are significant. Yet, some influence remains, so, for detailed analyses an incidence angle normalization might be applied to remove eventual biases in VV backscatter. Furthermore, this normalization should take into account the dynamic nature of crops and the variations in $\gamma^0 = f(\theta_{loc})$ relationship during the year.

4.3. Methodology

4.3.1. Backscatter sample selection

The objective of this analysis is to normalize γ^0 VV backscatter time series obtained in different orbits, so that they can be used in further analyses as a single time series. To evaluate the success of the normalization, a backscatter sample selection is done, selecting for each parcel acquisitions of different orbits obtained in a small time-frame. Ideally, in case simultaneous acquisitions were available, eventual biases between acquisitions would fade after a successful normalization. Longer time-frames enhance the probability of backscatter variations between orbits due to other 'disturbing' factors (e.g., precipitation). Orbits 8DESC and 103ASC overpassed the study area the same days at 6:00 and 18:00, respectively; and orbit 81DESC 24 hours before 8DESC, and thus 36 hours before 103ASC. Since significant changes in wheat conditions are not expected in 36 hours, backscatter values in this time-frame are considered comparable unless a strong weather event or agricultural practice (e.g., tillage or harvest) occurred in between. Thus, to mask out these eventual disturbing factors, for each parcel, dates with a γ^0 backscatter difference between orbit pairs larger than 3 dB were excluded from the analysis. This excluded only ~15% of the data. To summarize, for each wheat parcel a sample of backscatter triplets (in orbits 8DESC, 81DESC and 103ASC) acquired during the two agricultural seasons was extracted and this formed the basis for all subsequent analyses.

4.3.2. Incidence angle normalization

The goal of normalization techniques is to remove the contribution of the incidence angle to the total backscatter. Mladenova et al. (2013) proposed a technique based on a histogram matching procedure that can account for the nonlinear nature of backscatter – incidence angle relationship. In this study, this technique was applied individually for the different periods explained below. The large dataset used in this study allowed

calculating the statistics needed for the histogram normalization at the agricultural region scale, as follows:

$$\gamma^0(ref) = \bar{\gamma}^0_{ref} + \hat{\gamma}^0_{ref} \frac{(\gamma^0 - \bar{\gamma}^0)}{\hat{\gamma}^0} \quad (4.1)$$

where γ^0 is radar backscatter in [dB]; '–' and '∧' indicate mean and standard deviation for each 1º incidence angle bin; and *ref* refers to reference angle. The reference angle was selected as the median incidence angle of the three orbits. As a result, the 40° reference angle was chosen.

Considering the different backscatter dependence on incidence angle (Fig. 1), two periods were considered: May-June and the rest of the year.

4.3.3. Azimuthal anisotropy correction

After incidence angle normalization, backscatter time series acquired in ASC and DESC orbits were subsequently processed to correct their eventual differences due to their different observation directions (azimuthal anisotropy) following Schaufler et al. (2018) . If no azimuthal anisotropy existed, the orbits' means should be the same:

$$\bar{\gamma}^0(40^o)_{103ASC} = \bar{\gamma}^0(40^o)_{8DESC} = \bar{\gamma}^0(40^o)_{81DESC} \quad (4.2)$$

To achieve this, the azimuthal correction method (Schaufler et al., 2018) computes first a reference backscatter value as the mean of all backscatter data from the three orbits:

$$\gamma^0_{ref} = \bar{\gamma}^0(40^o) \quad (4.3)$$

Then, the difference between this reference value and the mean of each orbit is the correction factor d_{orbit} necessary to compensate azimuthal effects:

$$d_{orbit} = \gamma^0_{ref} - \bar{\gamma}^0(40^o)_{orbit} \quad (4.4)$$

The azimuthal anisotropy is finally corrected by adding d_{orbit} to each normalized backscatter value:

$$\gamma^0(40^o)_{az\_corr} = \gamma^0(40^o)_{orbit} + d_{orbit} \quad (4.5)$$

All the normalized backscatter data from section 3.2. were processed with this algorithm. The two periods were also separated for the anisotropy correction.

4.3.4. Evaluation of results

A successful correction of incidence angle and azimuthal anisotropy contributions would produce the same parcel scale backscatter values acquired in different orbits within the established 36 hour time-frames. Therefore, the absolute backscatter difference (|Δ|) from the three orbits was used as an evaluation criteria. The |Δ| between orbits were computed pairwise (8DESC-103ASC, 8DESC-81DESC and 81DESC-103ASC) for each parcel. These |Δ| values were grouped per month to illustrate the performance of the correction methods throughout the year.

Additionally, the performance of the corrections was evaluated using the correlation between backscatter and Red-edge NDVI (reNDVI) (Gitelson and Merzlyak, 1994). The rationale of this comparison is that if the applied corrections successfully reduced angular effects, then the correlation with reNDVI should improve. For this, a ~270 km^2 pilot zone was selected (N42º42'53.6'' and W1º16'12.7''), containing 855 wheat parcels with different topographic orientations and slopes. Sentinel-2 scenes were searched, but cloud affection is typically persistent during wheat growing cycle in Navarre. Therefore, the correlation was assessed for a complete agricultural campaign (2018), but also focusing only on its final part, i.e. April-July, where the availability of cloud-free scenes improves. This period includes crop maturity, senescence and harvest. In total, 14 cloud-free Sentinel-2 Level-2A scenes were obtained. The median reNDVI time series for each parcel were obtained and interpolated for Sentinel-1 acquisition dates. The Pearson correlation coefficient R was computed between the interpolated reNDVI time series and all the backscatter time series (σ^0, γ^0, $\gamma^0(40º)$ and $\gamma^0(40º)_{az\_corr}$).

4.4. Results and discussion

Figure 4.2 represents the |Δ| values for all wheat fields during the season and for the four different processing alternatives considered: σ^0, γ^0, $\gamma^0(40º)$ and $\gamma^0(40º)_{az\_corr}$. σ^0 backscatter had the highest |Δ| values, with a median of 1.25 dB that represented the significant influence of incidence angle variations in backscatter. γ^0 presented a median |Δ| of 0.80 dB, with October and June being the months with the lowest values (~0.67 dB). The significant reduction of |Δ| (~0.45 dB) for γ^0 (Fig. 2), illustrated the effectiveness of terrain flattening for compensating incidence angle variations in backscatter data, in accordance with Fig. 4.1.

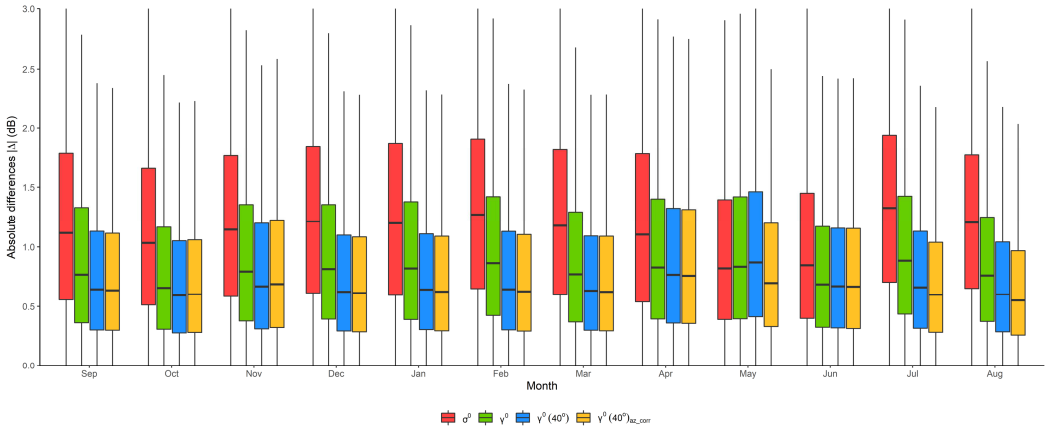

Figure 4.2. Boxplots of absolute backscatter differences for the three orbits during the year and considering different processing alternatives.

|Δ| decreased slightly further after incidence angle normalization ($\gamma^0(40º)$), with a reduction of 0.15 dB compared to σ^0, and also after azimuthal anisotropy correction ($\gamma^0(40º)_{az\_corr}$), where |Δ| achieved an additional reduction of 0.02 dB. Both corrections were effective but the reduction was much lower than the initial reduction achieved after terrain flattening. The intensity of the incidence angle normalization differed between months, and was highest for months with no vegetation, i.e., July, August; and winter months (BBCH growth stages 1-2), where the reductions of |Δ| were >0.16 dB. For May and June $\gamma^0(40º)$ had almost no effect in |Δ|, due to the reduced effect of the incidence angle in backscatter during these months (Fig. 1). Conversely, $\gamma^0(40º)_{az\_corr}$ was most effective in May, compensating the undercorrection of γ^0 and $\gamma^0(40º)$. In summary, the median |Δ| for all the months was 0.65 dB for $\gamma^0(40º)$, and 0.63 dB for $\gamma^0(40º)_{az\_corr}$.

When evaluating |Δ| values per month, it was possible to observe that the period after harvest (July and August) achieved the highest |Δ| reductions after both corrections (>0.21 dB). In September, October and November, corrections had a smaller effect (~0.10 dB). During these months sowing occurs, leading to a higher soil roughness variability between different parcels that might mask potential improvements for individual parcels. Tilled or newly sown soils have a more Lambertian behavior, leading to a lower influence of incidence angle (Ulaby et al., 1982). December, January and February were the months that presented the highest |Δ| in the original data and achieved a good reduction of differences after incidence angle normalization (~0.20 dB). In March and April no significant improvements were detected (<0.14 dB). This could be explained by the dynamics of VV backscatter during

wheat growth cycle. Once wheat initiates the stem elongation stage, an attenuation of VV backscatter occurs (Mattia et al., 2003), leading to lower differences between subsequent acquisitions. May and June were the months with the lowest influence of incidence angle. While the anisotropy corrections achieved a certain decrease of |Δ| in May (0.13 dB), there was little improvement in June (0.02 dB).

Exploring the influence of terrain slope and orientation on the mean |Δ| of each wheat parcel (Figure 4.3) it can be observed that terrain flattening successfully reduced the majority of the radiometric effect of topography. Yet a residual dependence remained, which could be further reduced with the incidence angle and azimuth anisotropy corrections.

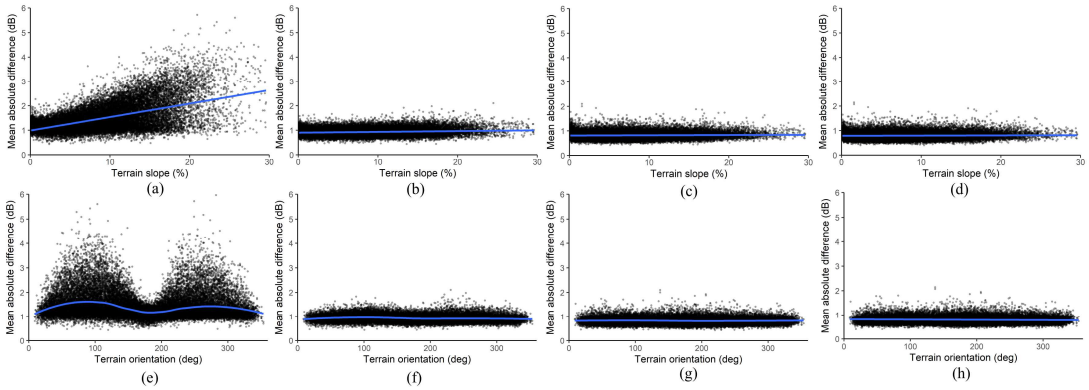

Figure 4.3. Dependence of mean absolute backscatter difference of wheat parcels on terrain slope (upper row) and orientation (lower row); for the different backscatter outputs investigated: (a, e) σ^0, (b, f) γ^0, (c, g) $\gamma^0(40º)$, (d, h) $\gamma^0(40º)_{az\_corr}$.

The correlation results for reNDVI and backscatter (Table 4.1) showed a negative correlation that responded to the typical backscatter pattern of wheat at VV polarization (Mattia et al., 2003). Correlation values for the full campaign were rather low, but when focusing on the period between maturity and harvest it improved as a consequence of the rising backscatter values at this period (Brown et al., 2003; Veloso et al., 2017) and the rapid decrease of reNDVI at crop ripening and senescence (Fig. 4.4). The improvements in correlation were minor for the full campaign (from 0.40 to 0.42), but more significant from maturity to harvest (from 0.59 to 0.69). These results demonstrate that the corrections applied (in particular γ^0 and $\gamma^0(40º)$) enhance the correlation with optical vegetation indices, and hence provide a better description of wheat growth. Yet, further studies in areas with a higher availability of optical data should be performed to confirm these results.

Table 4.1. Pearson correlation coefficient of reNDVI and backscatter for the pilot zone

Backscatter correction	Full campaign	Maturity to harvest
σ^0	-0.396	-0.587
γ^0	-0.417	-0.662
$\gamma^0(40º)$	-0.419	-0.686
$\gamma^0(40º)_{az\_corr}$	-0.419	-0.686

The backscatter time series (Fig. 4.4) showed that the corrections were successful as they not only improved the matching of the three orbits but also reduced backscatter variability (error bars), since their eventual differences due to angular effects also decreased.

Taking into account that Sentinel-1 radiometric accuracy is 1 dB (Berger et al., 2012), for some applications terrain flattened backscatter (γ^0) computed for different orbits might be comparable without further processing given the relatively low influence of acquisition geometry in the values observed here. However, for quantitative analyses requiring more detail (e.g. soil moisture retrieval) or when different orbits need to be combined to enhance the temporal resolution, a closer match between these orbits might be achieved by implementing incidence angle normalization and azimuthal anisotropy correction techniques (Bauer-Marschallinger et al., 2021). Our results, validate the correction methods applied (Mladenova et al., 2013; Schaufler et al., 2018; Small, 2011) and recommend their implementation in image processing pipelines and software. Coinciding with our results, recently (d'Andrimont et al., 2021), the importance of terrain flattening and incidence and azimuth angle corrections for operational applications of Sentinel-1 data was stressed out, in particular for crops with prolonged bare soil phases.

Although a general idea is that agricultural lands occupy flat terrains, in many parts of the world, farmers cultivate areas that have significant slopes, making these corrections necessary even when working with a single orbit, as our results confirm. In particular, it is shown that for wheat the incidence angle influence on backscatter changes during the season due to its phenological development. Similar studies in other crops are recommended to confirm this finding.

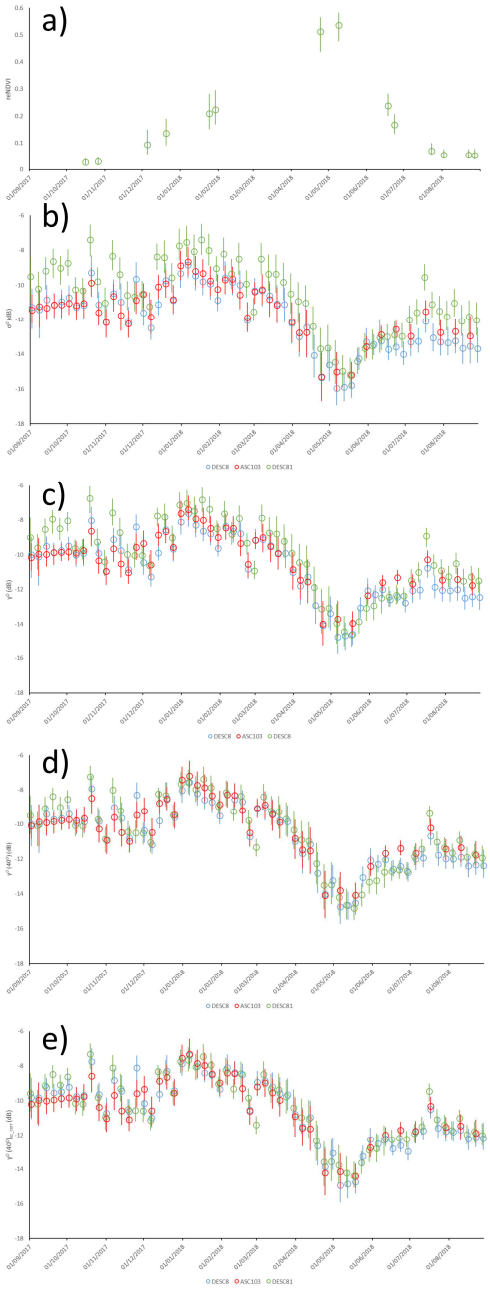

Figure 4.4. Median time series of reNDVI and backscatter for wheat parcels in the pilot zone. The error bars represent the interquartile range (IQR): (a) reNDVI (b) σ^0, (c) γ^0, (d) $\gamma^0(40º)$, (e) $\gamma^0(40º)$az_corr. Colors in b, c, d and e represent the different orbits.

4.5. Conclusion

In this study, the influence of acquisition geometry (incidence and azimuth angles) on backscatter (σ^0) and terrain flattened backscatter (γ^0) was evaluated for wheat parcels. The analysis revealed that terrain flattening markedly reduced the influence of incidence angle in VH and VV polarizations, being almost negligible for VH polarization. In VV polarization, the influence of the incidence angle slightly remained, although it varied along the growing season, being the least when the crop canopy was fully grown. It was thus demonstrated that the incidence angle influence on backscatter varied due to the phenological development of the crop.

The analysis of backscatter differences between the three relative orbits studied, showed that terrain flattening could achieve a significant reduction of angle variations in backscatter data. Yet, incidence angle normalization could further reduce backscatter differences, particularly in winter months and bare soil periods. The correlation with reNDVI also improved after terrain flattening and normalization, revealing a better description of wheat growth. The azimuth anisotropy correction had a lower effect that was mainly relevant in May. The differences between ascending and descending passes might be partly due to the acquisition geometry (incidence and azimuth angles) and partly due to the time of the day (eventually dew, frost, soil moisture, etc.), but this has not been sufficiently studied yet.

For applications where different orbits have to be combined, the need to further correct terrain-flattened backscatter values will depend on the level of precision required. Furthermore, for quantitative studies aiming at retrieving a bio-geophysical variable of interest (e.g., soil moisture), adding these corrections might provide enhanced results.

CHAPTER 5

A NEW METHODOLOGY FOR WHEAT ATTENUATION CORRECTION AT VV-POLARIZED BACKSCATTER TIME SERIES

Published in: Arias, M., Campo-Bescós, M.A., Arregui, L.M. González-Audícana, M. and Álvarez-Mozos J. 2022. A new methodology for wheat attenuation correction at C-band VV-polarized backscatter time series. IEEE Transactions on Geoscience and Remote Sensing, Vol. 60. DOI: 10.1109/TGRS.2022.3176144.

Supplementary materials available at: https://ieeexplore.ieee.org/document/9777738

Abstract

Wheat is one of the most important crops worldwide, and thus the use of remote sensing data for wheat monitoring has attracted much interest. Synthetic Aperture Radar (SAR) observations show that, at C-band and VV polarization, wheat canopy attenuates the surface scattering component from the underlying soil during a significant part of its growth cycle. This behavior needs to be accounted for or corrected before soil moisture retrieval is attempted. The objective of this paper is to develop a new method for wheat attenuation correction (WATCOR) applicable to Sentinel-1 VV time series and based solely on the information contained in the time series itself. The hypothesis of WATCOR is that without attenuation, VV backscatter would follow a stable long-term trend during the agricultural season, with short-term variations caused by soil moisture dynamics. The method relies on time series smoothing and changing point detection, and its implementation follows a series of simple steps. The performance of the method was compared by evaluating the correlation between backscatter and soil moisture content in six wheat fields with available soil moisture data. The Water Cloud Model (WCM) was also applied as a benchmark. The results showed that WATCOR successfully removed the attenuation in the time series, and achieved the highest correlation with soil moisture, improving markedly the correlation of the original backscatter. WATCOR can be easily implemented, as it does not require parameterization or any external data, only an approximate indication of the period where attenuation is likely to occur.

Keywords: wheat, attenuation, SAR, time series, soil moisture, Sentinel-1

5.1. Introduction

Wheat is one of the most important crops at the global scale, being the main food grain source for humans (FAO, 2021). The monitoring of major crops like wheat is essential for important applications such as food security assurance (Van Tricht et al., 2018) or biomass and yield forecasting (Steele-Dunne et al., 2017). Soil moisture (SM) is one of the key variables subject to be monitored. The occurrence of plagues and diseases can be affected by SM, and it plays a key role in the development of crops as it determines the availability of water for plants (Vereecken et al., 2010). At a global scale, SM is a key variable of the climate system and it is involved in a number of feedbacks affecting weather events (Seneviratne et al., 2010).

In-situ probes can measure SM at the point scale, but generalizing point measurements to spatial areas is not straightforward due to its high spatial variability (Ochsner et al., 2013). Therefore, remote sensing has received great interest, as a source of spatial information over large areas of the territory with a given periodicity. Microwave sensors are the most suitable for SM estimation, because at this wavelengths the soil response is determined partly by its dielectric properties that mainly depend on surface SM (Kornelsen and Coulibaly, 2013). Operational global SM products at coarse resolution (25-50 km) have been developed in the last years, mainly based on radiometers (Brocca et al., 2011; Chan et al., 2016; Kerr et al., 2012) or scatterometers (Naeimi et al., 2009). SAR sensors offer a finer spatial resolution (10-20 m), suitable for SM estimation at the agricultural field or irrigation sector scale. Nevertheless, these sensors are more severely affected by other variables such as soil surface roughness (Verhoest et al., 2008) or vegetation characteristics (Bindlish and Barros, 2001), and thus, SM retrieval at the field scale is still a challenging task (Peng et al., 2021; Wagner et al., 2007).

SAR based SM retrieval is a subject that has been studied for more than forty years (Macdonald and Waite, 1971). Different models have been developed for bare soils over the years. The Integral Equation Model (IEM) (Fung, 1994) and the Advanced Integral Equation Model (AIEM) (Chen et al., 2015; Fung and Chen, 2004) are physical-based models widely used (Kong et al., 2018; Shi et al., 1997). Semi-empirical models were also developed, such as Oh (Oh et al., 1992), Dubois (Dubois et al., 1995) or Shi (Shi et al., 1997). The inversion of these models allows estimating a variable (e.g. SM) from backscatter observations, knowing the rest of the variables of the model (e.g., surface roughness). Machine learning techniques have also been used for SM estimation, for instance artificial neural networks (ANN) (Hachani et al., 2019) or support vector machines (SVM) (Ahmad et al., 2010; Pasolli et al., 2015a). In some studies, machine learning approaches were trained using synthetic datasets generated with models, such

as IEM (Hajj et al., 2017), while in other cases training was performed with real remote sensing observations and in situ data (Greifeneder et al., 2021).

The backscatter response from vegetation canopies is complex, as it is influenced by sensor configuration (e.g., frequency, wavelength and incidence angle), the physical structure and the dielectric properties of plant elements and the characteristics of the underlying soil. Therefore, SM estimation under vegetated surfaces requires the coupling of vegetation and soil backscatter models (Zhang et al., 2021) in order to separate both contributions. Different models that simulate the backscatter from canopies have been developed. Full electromagnetic scattering models, like the Michigan Microwave Canopy Scattering model (MIMICS) (Brown et al., 1994; Ulaby et al., 1990), might be difficult to use in an operational setting due to the large number of parameters required. Therefore, approximate solutions like the semi-empirical Water Cloud Model (WCM) (Attema and Ulaby, 1978) have gained interest and popularity due to its relatively simplicity (Graham and Harris, 2003; Kornelsen and Coulibaly, 2013). The WCM represents the vegetation canopy as a medium composed of identical water particles that can be represented by bulk descriptors related to its density (e.g., Vegetation Water Content or LAI). For empirically fitting WCM coefficients, backscatter observations, SM measurements and vegetation descriptors are required. Different studies investigated alternative vegetation descriptors that might be obtained from in-situ data or from remote sensing observations (Bindlish and Barros, 2001; Hajj et al., 2017). The first ones are costly and time consuming, while remote sensing based descriptors might be more easily obtained. The latter include optical vegetation indices (e.g., NDVI.) (Bao et al., 2018; Zhuo et al., 2019) or features obtained from multi-pol SAR observations (crosspol ratio, etc.) (Dabrowska-Zielinska et al., 2018; Li and Wang, 2018). SAR features might be particularly interesting for areas where weather conditions preclude the use of optical data (Li and Wang, 2018). WCM simulates both the vegetation contribution to backscatter (volume scattering) and its effect in attenuating the soil contribution. It usually requires a specific parameterization for each vegetation type, vegetation descriptor and study site (Álvarez-Mozos et al., 2006), which needs additional in situ-measurements for validation purposes.

Wheat canopy has a particular behavior in VV polarization, with a characteristic scattering dynamics during the different phenological stages (Stiles et al., 2000), and a dominant scattering mechanism consisting of soil backscatter attenuated by the canopy (Brown et al., 2003; Mattia et al., 2003; Skriver et al., 1999b). At early stages backscatter mainly reflects the response of soil, since the crop development is still poor (Harfenmeister et al., 2019; Larranaga et al., 2013). Then, this soil backscatter component is attenuated gradually during the stem elongation phase due to the vertical structure of wheat plants, reaching its minimum values by the heading stage (Mattia et

al., 2003), when backscatter starts to increase. Then, during crop maturity and senescence, wheat plants dry out and the attenuation capacity of wheat canopy gradually diminishes leading to an increase in backscatter (Liu et al., 2013b) until the crop is finally harvested. Ouaadi et al (Ouaadi et al., 2021) proposed that the attenuation produced by wheat was the main cause of the scattered relationship between SM and backscatter, while Weiβ et al (Weiß et al., 2021) found that the uncertainty in SM estimation increased from stem elongation until heading, coinciding with the period of attenuation.

An approach for attenuation correction that would not require external data or local-site parameterization would be of wide interest from the applications point of view. Detailed backscatter time series observed over wheat fields by ongoing missions with a systematic observation planning, like Sentinel-1, might be an invaluable information source for characterizing the attenuation behavior of wheat. This would allow designing and implementing data-driven approaches based on the backscatter time series itself. Therefore, the objectives of this paper are:

- To describe the backscattering behavior of wheat for C-band VV polarization using a large dataset of Sentinel-1 observations.

- To propose a new method for wheat attenuation correction (WATCOR) in VV polarization based solely on the backscatter time series itself.

- To evaluate the proposed approach and compare it with the state-of-the-art WCM using a dataset comprising in-situ SM measurements over several wheat fields.

The remaining of the article is structured in three separate sections, section 5.2 describes the general backscatter behavior of wheat, section 5.3 proposes the new method WATCOR and section 5.4 evaluates and compares it with the WCM. Finally some conclusions are drawn

5.2. General wheat behavior at C-band VV backscatter

This section analyzes a large wheat backscatter dataset collected during four agricultural campaigns (2016, 2017, 2018 and 2019) in Navarre (Spain). The objective is to describe the general backscatter behavior of this crop in VV polarization.

5.2.1. Methodology

a) Study area

The study area comprises the agricultural lands of the province of Navarre in Northern Spain (Figure *5.*1). In spite of its relatively small extension (10,391 km^2), Navarre is characterized by its diversity regarding climate and landscape. The Northern area corresponds with the western side of the Pyrenees mountain range, and has a humid climate with a predominance of forests and prairies. In contrast, the Southern area, with a drier climate and a higher proportion of arable land, is formed by the plains of the Ebro basin. The transition zone between these two areas has mixed characteristics. This diversity leads to a stratification of Navarre in seven agricultural regions for administrative and management purposes. The detailed characteristics of the agricultural regions are described in Arias et al. (Arias et al., 2020). Table 5.1 presents some basic features of the wheat fields therein.

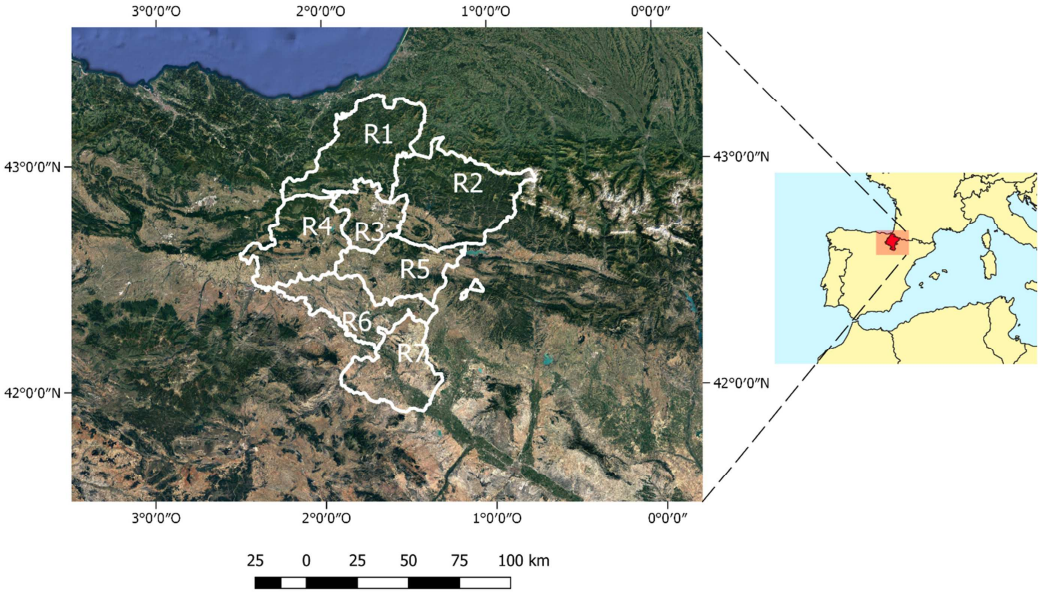

Figure 5.1. Location of the province of Navarre and its seven agricultural regions

Table 5.1. Characteristics of wheat fields on each agricultural region

Agricultural region	Wheat area cultivated (ha)	Average size of fields (ha)	Average slope of fields (%)	Type of management
R1	583	1.03	12	Rainfed
R2	8126	1.83	11	Rainfed
R3	11620	1.73	11	Rainfed
R4	16990	1.39	11	Rainfed
R5	14941	1.47	9	Rainfed/Irrigated
R6	10348	1.55	5	Rainfed/Irrigated
R7	7695	1.73	6	Rainfed/Irrigated

b) SAR imagery

Sentinel-1 C-band SAR imagery was the base for this study. Images were acquired in the Interferometric Wide (IW) swath mode with dual-pol (VH-VV) configuration, and they were downloaded as level-1 Ground Range Detected (GRD) products. Sentinel-1 overpasses the study area in one ascending orbit (103ASC) and two descending orbits (8DESC and 81DESC). All available Sentinel-1A and B scenes in these three orbits covering the study area from 1 September 2015 to 31 August 2019 were used in this study, making a total of 563 scenes. The revisit time was variable during the years of study and the different orbits used. In 2016, S1B was not yet available so the revisit time was 12 days in all orbits, and then it shortened to 6 days for the rest of the study, except for orbit ASC103 where S1A was unavailable from April 2018 to the end of 2019.

The images were processed with an automated pipeline in SNAP Graph Processing Toolbox following these steps: 1) thermal noise removal; 2) slice assembly; 3) apply orbit file; 4) calibration; 5) speckle filtering (3x3 Gamma-Map); 6) terrain flattening; 7) range-doppler terrain correction and 8) subset to the extent of Navarre. After the process, γ^0 backscatter coefficients in dB units were obtained. The pixel size of the output products was set to 20 m. The SRTM 1sec HGT DEM was used for terrain flattening and terrain correction.

c) Data extraction

The Agriculture Department of the Government of Navarre provided an anonymized version of the Land Parcel Information System (LPIS or SIGPAC in Spanish) with the crops declared for each parcel in the 2016, 2017, 2018 and 2019. This dataset consisted of a GIS polygon vector file and its corresponding attribute table with the crops stated by farmers in their EU Common Agricultural Policy (CAP) declarations. From this dataset, wheat parcels were extracted for each year and preprocessed as follows: (1) parcels smaller than 0.5 ha were discarded, (2) a 5 m inner buffer was applied to the

parcel boundaries, (3) the median backscatter was computed for each parcel in each acquisition date leading to a backscatter time series for each parcel and year, and (4) the 10% of parcels most dissimilar to the characteristic wheat time series were discarded each year, since some CAP declarations might be erroneous (Arias et al., 2020). The final number of wheat parcels used in this study was 21845 for 2016, 18750 for 2017, 20374 for 2018 and 20465 for 2019.

d) Analysis of backscatter trend

For each agricultural region, year and satellite pass, the median VV backscatter time series of all wheat parcels were computed. Additionally, the median time series at the province level (Navarre) were also calculated. Each agricultural year started the 1st September of the previous year and ended the 31st August of the corresponding year.

Before analyzing the backscatter behavior of wheat, the similarity between years was evaluated by comparing the autocorrelation function (ACF) of the median time series. The ACF measures how fast or slowly data in the time series vary (Broersen, 2006), and it is useful for identifying underlying trends in the time series. The correlation length (l_{cor}) was also calculated and used as a parameter that summarizes the information conveyed in the ACF. l_{cor} is the lag distance (days in our case) where autocorrelation in the time series is lost, this is commonly defined as the lag where the ACF falls below $1/e$ (Ulaby et al., 1982). Finally, in order to evaluate the long-term backscatter trend, a moving average of 36 days was computed for the median time series. This process enables omitting rapid backscatter variations due to meteorological events or eventual agricultural practices, so that only the general backscatter trend caused by wheat canopy remains in the time series.

5.2.2. Autocorrelation results

Figure 5.2 showed that the ACFs obtained for the different years were similar, particularly during the first 50 lagged days. Autocorrelation values steadily decreased to a value of zero around day 100. Additionally, the ACF between the different orbit nodes was also quite similar, although orbit 103ASC had a slightly steeper decay.

When computing the ACFs for each agricultural region (Supplementary Materials 1), it can be observed that regions 6 and 7 presented a steeper decay than the rest, meaning that backscatter variations in wheat parcels of these regions might be more abrupt than in the other regions. The correlation length (Table 5.2) enables comparing these ACFs quantitatively. Although there was some variability in l_{corr} between the different years studied, values did not deviate drastically from the mean (~30 days). The differences observed could be attributed as a first instance to the climatic conditions, with 2018

being the wettest year in the series for regions 4, 6 and 7 and showing lower l_{corr} values than the remaining years. In addition, Sentinel-1 revisit time was not consistent during the whole study period. It concurred that time series with lower temporal resolution (i.e., longer revisit time) had lower l_{corr}. Agricultural regions 6 and 7 showed lower l_{corr} values, probably due to the larger amount of irrigated wheat parcels in these regions that might result in a more dynamic SM variation (Supplementary Materials 1). The remaining regions had a very similar autocorrelation behavior to that observed at the provincial level. Altogether, this analysis demonstrates that wheat backscatter time series have a certain degree of autocorrelation that is similar every year, suggesting that wheat canopy creates an annual systematic trend in VV backscatter.

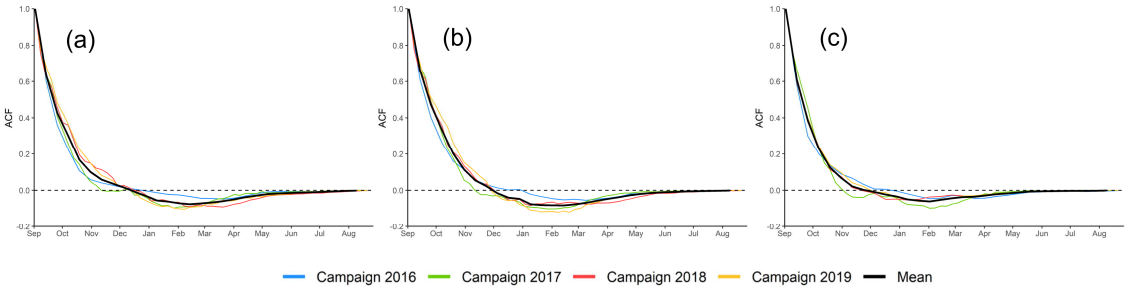

Figure 5.2. Autocorrelation function (ACF) plots of median wheat time series in Navarre for the three different orbits: (a) 8DESC, (b) 81DESC and (c) 103ASC

Table 5.2. Correlation length of median wheat time series ACF

Orbit node	Campaign 2016 (days)	Campaign 2017 (days)	Campaign 2018 (days)	Campaign 2019 (days)	Mean (days)
8DESC	25	28	34	37	31
81DESC	28	33	36	41	35
103ASC	22	29	27	27	26

5.2.3. Long-term wheat backscatter trend analysis

Figure 5.3 showed that backscatter slightly increased from September to November due to soil preparation and sowing. Also, November is normally a wet month, so the backscatter rise might respond to higher soil moisture. During December and January, backscatter values remain rather high (~-9.5dB). However, there was a clear decrease in backscatter from February to April, which reached its minimum value (~15dB) at the end of April. During this period, the vertical orientation of wheat canopy at the stem elongation stage (BBCH21-55) (Lanchashire et al., 1991) produces an attenuation of backscatter (Mattia et al., 2003). Then, backscatter increased from April to June due to changes in the structure of plants, as a consequence of the successive phenological development, from flowering to ripening (BBCH55-99). Finally, after harvest, a

stabilization in backscatter values was found in July and August (~12dB), when backscatter depends only on the bare soil response. Although September is bare soil month, it presents higher backscatter values compared to August due to the increase of soil roughness caused by soil preparation.

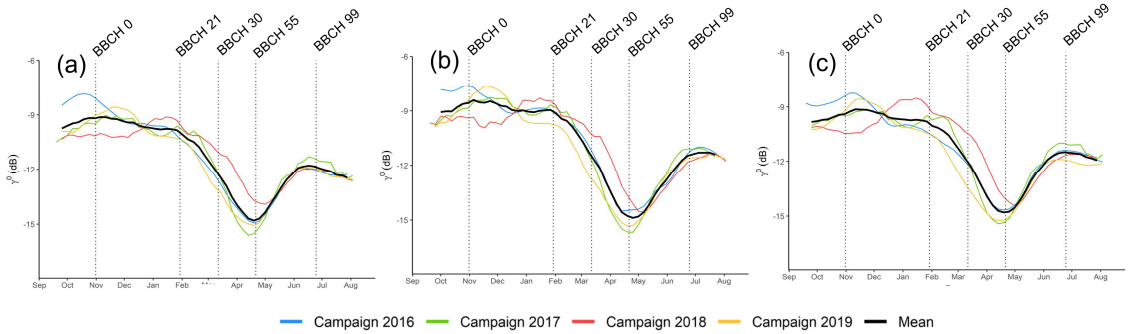

Figure 5.3. Smoothed Sentinel-1 VV time series for the different agricultural campaigns and orbits: (a) 8DESC. (b) 81DESC. (c) 103ASC. Curves represent the median of all wheat parcels in Navarre. Main phenological stages (BBCH scale) are indicated.

The smoothed time series of the different agricultural regions (Supplementary Materials 2) exhibited the same behavior with differences in their backscatter amplitude. For instance, the decrease in backscatter during attenuation reached ~6dB in northern regions, while in southern regions it was lower (~4dB). The duration of the attenuation period also varied, being longer in southern regions.

In summary, the analysis of the smoothed backscatter time series confirmed the existence of a seasonal trend caused by wheat canopy in VV backscatter, rather independent from SM variations and explained by wheat growth and phenology. Backscatter attenuation caused by the growth of vertical stems started approximately in February and lasted until the end of April, when the maximum attenuation was observed. Then, backscatter increased due to phenological development and later also due to the greater penetration when cereals are drying out during ripening and senescence, until harvest took place at the end of June. For individual fields, both the start and end of attenuation might be identified as changing points in their backscatter time series. The specific start and end points of attenuation for each parcel might vary due to the particular management and site conditions of the parcel.

5.3. A new method for correcting wheat attenuation in VV backscatter time series

The previous section demonstrated that wheat canopy produces a particular backscatter pattern in VV polarization during the agricultural year. This section

proposes a simple methodology for correcting wheat attenuation in VV backscatter time series (WATCOR), by using only information contained in the VV backscatter time series itself.

The initial hypothesis is that if no attenuation existed, VV backscatter would follow a rather stable long-term trend with backscatter variations mainly caused by surface roughness or moisture variations, as it occurs in the initial (Sep-Jan) and final (Jun-Aug) bare soil periods.

The method follows a series of steps (Figure 5.4). First, the period where attenuation occurs is identified. Then, a low frequency trend is interpolated for this period by assuming a linear transition from the start of the attenuation period to its end. Finally, the high frequency backscatter variations are finally added to the low frequency trend. These high frequency variations in the attenuation period are extracted by fitting and subtracting a lower envelope curve to the real backscatter time series. In the following subsections, these processes are explained in detail.

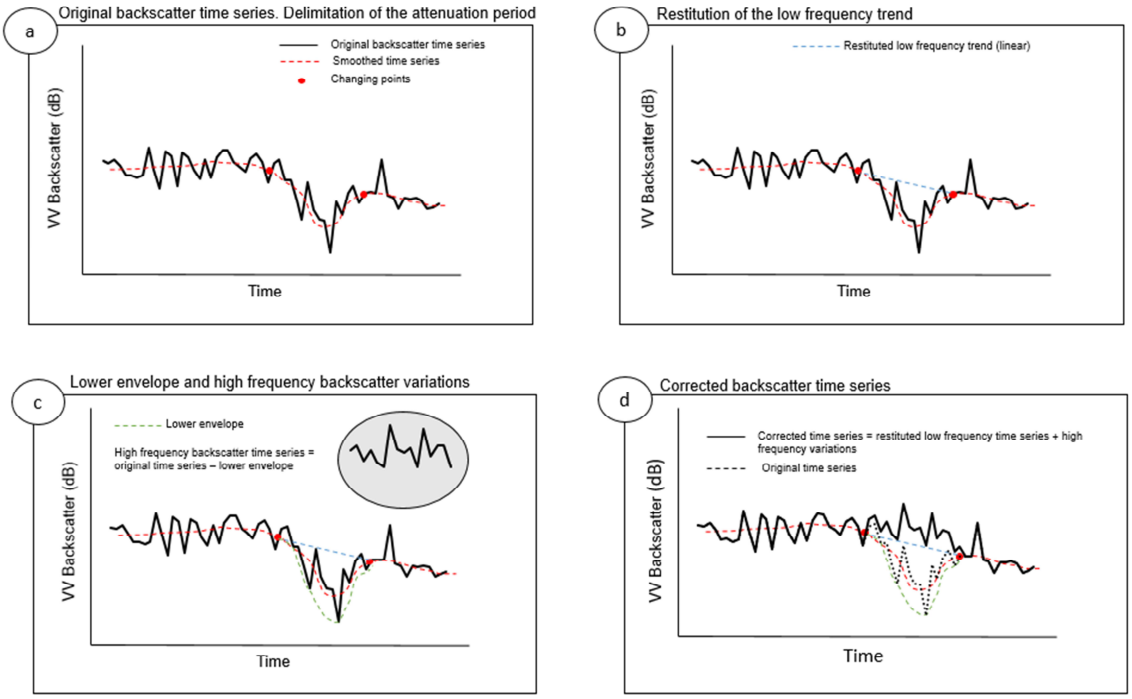

Figure 5.4. Diagram indicating the main steps of WATCOR

5.3.1. Low frequency backscatter trend extraction

The low frequency trend underlying VV backscatter time series was extracted by applying a Savitzky-Golay smoothing filter (Liu et al., 2013a). This filter considers that high frequency variations are noisy, and obtains a smoothed trend by computing a polynomial least-squares fit of order *p* inside a moving window of width *w* that crosses the signal (Savitzky and Golay, 1964). The Savitzky-Golay filter requires a signal with constant frequency. Therefore, to comply with the missing scenes that seldom occur in Sentinel-1 time series, a daily linear interpolation was performed first. Then, the parameters of the filter were determined through visual interpretation. High values of *w* and low values of *p* produce smoother trends. The effects of varying these two parameters are displayed in Supplementary Materials 3. A polynomial order of one (p=1) with a window size of 45 days (w=45) yielded the best results, obtaining a smooth signal that preserved the attenuation trend and omitted high frequency backscatter dynamics.

5.3.2. Identification of the attenuation trend

The period when wheat attenuated VV backscatter was identified by applying a change point detection technique in the smoothed backscatter series. Change points are points in the time series where unexpected and structural changes occur, changing the data properties such as the mean or the variance (Militino et al., 2020; Sharma et al., 2016). Both the start and end of the attenuation period are changing points in the smoothed trend. Although the Savitzky-Golay smoothing minimized the number of changing points detected outside the attenuation period, some additional change points might be found, depending on other factors, such as the climatic conditions of the year. To avoid this, the change point search was constrained in time to periods where the start and end of attenuation are likely to happen. In our case, based on local knowledge, and the very high number of wheat parcels analyzed, the search was constrained to a 2-month period:

- Beginning of attenuation: from 15/Jan to 15/Mar.
- End of attenuation: from 15/May to 15/Jul.

These dates should be adjusted to the particular agricultural calendar and conditions of each site.

Once the search periods were set, a change point analysis algorithm (Matteson and James, 2014) was applied, based on the statistical moments of the time series. The algorithm performed a nonparametric estimation of the number of change points and their position based on a divisive hierarchical clustering, without any additional assumption on their distribution. The algorithm parameters where set to k=1 (number

of change points to estimate) and *min.size*=2 (minimum number of observations between change points).

5.3.3. Restitution of the smoothed trend

The smoothed backscatter trend was restituted to remove wheat attenuation. To achieve this, a linear trend was assumed between the start and end of the attenuation period (Figure 5.4b). Ideally, this restituted trend would follow the backscatter dynamics in case no attenuation existed.

5.3.4. Lower envelope determination

During the attenuation period, the lower envelope of the original backscatter time series represents the backscatter bottom line, with values above it that could correspond to soil wetting events. To extract this lower envelope, the iterative methodology of Chen et al. (Chen et al., 2004) was adapted to obtain a curve that fitted as good as possible the local minima in the backscatter time series. To automate the methodology six steps were followed (Figure 5.5):

1) Select the set of local minimum points in the backscatter time series (t, γ^0), through its first derivative. These points were flagged for later use as a reference $(t, \gamma^0)_{ref}$.
2) Smooth the time series using the Savitzky-Golay filter (parameters *p*=1, *w*=45), leading to a smoothed curve (t, γ^0_{s0}).
3) Initialization of an iteration (k = 1, … , 100). Remove high points in the time series, by selecting points that were above the smoothed curve and replacing their values with the ones on the smoothed curve (Eq.1). This way, a new backscatter time series was generated (t, γ^0_k),

$$\gamma^0_k = \begin{cases} \gamma^0, & \gamma^0 \leq \gamma^0_{sk-1} \\ \gamma^0_{sk-1}, & \gamma^0 > \gamma^0_{sk-1} \end{cases} \qquad (5.1)$$

4) Smooth again the obtained backscatter time series (t, γ^0_k) using the Savitzky-Golay filter (parameters *p*=2, *w*=45), producing a newly smoothed backscatter time (t, γ^0_{sk}), which got closer to the lower envelope.
5) Compute a fitting index (F_k), as the RMSE between the reference points $(t, \gamma^0)_{ref}$ and their values in (t, γ^0_{sk}).
6) Steps 3-5 where iterated 100 times, and finally, the lower envelope was selected as the iteration (t, γ^0_{sk}) with the lowest F_k

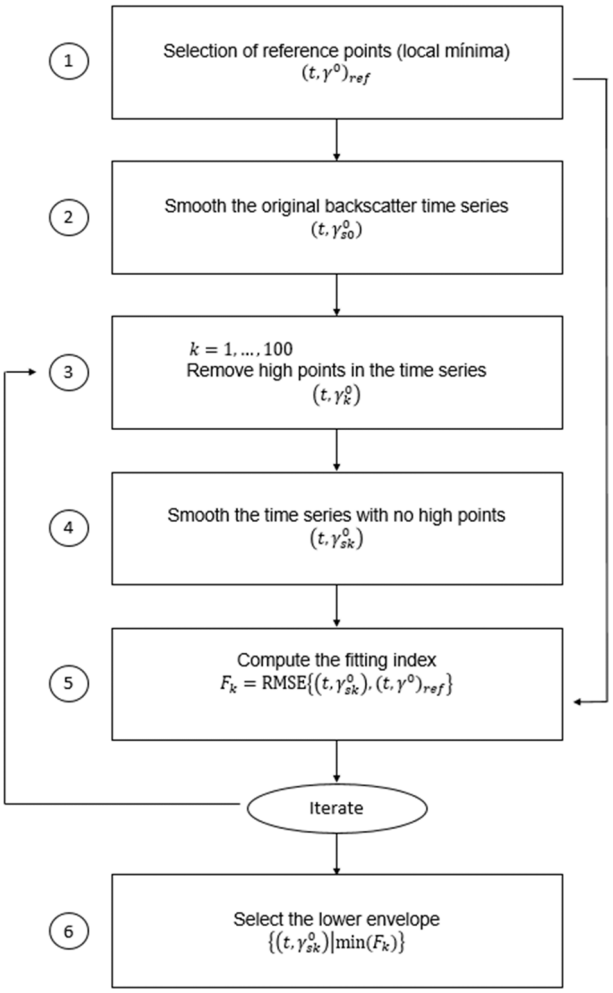

Figure 5.5. Flowchart explaining the determination of the lower envelope

5.3.5. Addition of high frequency backscatter variations to the restituted trend

The difference between the original backscatter time series and the lower envelope was calculated. This difference represented high frequency backscatter variations that mostly responded to soil wetting events. Therefore, it was added to the smoothed linear trend obtained in section C, resulting in the corrected backscatter time series, where wheat attenuation had been finally removed (Figure. 5.4d).

5.4. Case study: backscatter and soil moisture correlation

The aim of this last section is to evaluate the effectivity of the proposed approach (WATCOR) with a case study. Since the correction of the vegetation effect is a pre-requisite for many SM retrieval methods (e.g., IEM, Oh model, etc.), this evaluation was performed indirectly by measuring the correlation between SM measurements and VV backscatter time series before and after correction. With this aim, surface SM ground measurements obtained over wheat fields in three dedicated field campaigns were used. In addition, the proposed approach was compared with the well-known Water-Cloud Model (WCM) (Attema and Ulaby, 1978), implemented in three different variants. The rationale was that the method that best corrected vegetation influence should result in higher correlation values between backscatter and SM time series.

5.4.1. Methodology

a) Study area and soil moisture measurements

The study was conducted in six wheat fields located in the province of Navarre (Spain) (Figure 5.6). The main characteristics of the fields are described in table 5.3. From three to five capacitance SM probes (Sentek-multi) were installed on each field from winter until de end of June (harvest). The probes recorded the volumetric SM on a 30 minutes basis, at six different depths every 10cm, from the surface to 60cm deep, yet for this analysis only the surface layer (0-10cm) was considered. For each field, its SM time series was calculated as the median time series of the probes installed on it. Finally, from these field time series, a dataset was extracted with the SM measurements that coincided with Sentinel-1 acquisitions (Table 5.4.).

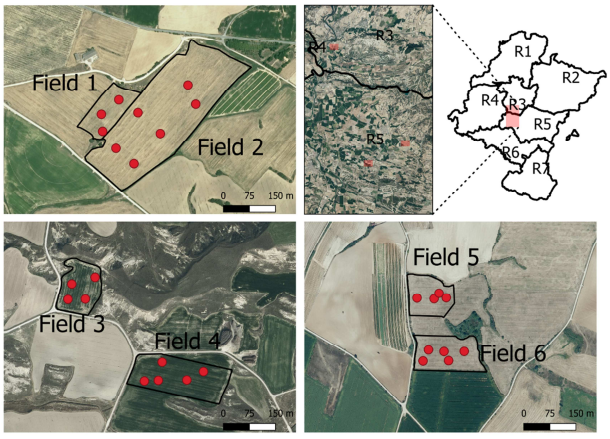

Figure 5.6. Wheat fields location. Red dots represent the soil moisture probes

Table 5.3. Characteristics of the wheat fields

Id	Area (ha)	Texture class	Type	Sowing date	Harvest	Number of probes
1	1.36	Clay-loam	Rainfed	25/10/2017	05/07/2018	3
2	8.50	Loam	Irrigated	25/10/2017	05/07/2018	6
3	1.57	Clay-loam	Rainfed	25/10/2018	06/07/2019	4
4	2.80	Silt-loam	Irrigated	25/10/2018	06/07/2019	5
5	1.37	Clay	Rainfed	11/12/2019	03/07/2020	4
6	1.86	Loam	Irrigated	11/12/2019	03/07/2020	5

Table 5.4. Number of Sentinel-1 acquisitions per orbit, and start and end dates of SM measurements

Id	8DESC	81DESC	103ASC	Initial date	Final date
1-2	20	21	13	03/03/2018	27/06/2018
3-4	19	20	9	14/02/2019	21/06/2019
5-6	29	29	29	31/12/2019	28/06/2020

b) Satellite imagery and data extraction

All available Sentinel-1 scenes from 1 September 2017 to 31 August 2020 were acquired and processed following the process explained in Section-II. After scene processing, a 10 m inner buffer was applied to the field boundaries and the median backscatter value for each field was computed on each of the Sentinel-1 acquisitions during its particular agricultural year (2018 for fields 1 and 2; 2019 for fields 3 and 4; and 2020 for fields 5 and 6). All backscatter data were normalized to a local incidence angle θ=40º, following the methodology explained in (Mladenova et al., 2013).

All available Sentinel-2 multispectral Level-2A Bottom of Atmosphere (BOA) reflectance images for the study period were also used, in this case obtained from Google Earth Engine. A subset of the study area was clipped, and scenes where the study fields were free of clouds were downloaded for further analysis. Two vegetation indices were computed: the normalized vegetation difference index (NDVI) (Rouse et al., 1974) and the normalized water difference index (NDWI) (Gao, 1996). As for Sentinel-1, median values were computed for each field, using the buffered polygon vector files, leading to field time series of NDVI and NDWI. Finally, a linear interpolation was applied to the NDVI and NDWI time series at a daily time step to derive a time series coincident with Sentinel-1 acquisitions.

c) WCM vegetation correction

The WCM is a semi-empirical model used to correct the vegetation influence on backscatter data (Attema and Ulaby, 1978). The model describes vegetation as a 'cloud' composed of identical water droplets uniformly distributed within the canopy, which attenuates the microwave radiation, but might also contribute to the total backscatter (Attema and Ulaby, 1978). The WCM is expressed as follows:

$$\sigma^0_{can} = \sigma^0_{veg} + \tau^2 \sigma^0_{soil} \quad (5.2)$$

$$\tau^2 = \exp\left(-\frac{2BV_2}{cos_\theta}\right) \quad (5.3)$$

$$\sigma^0_{veg} = AV_1 cos\theta (1 - \tau^2) \quad (5.4)$$

where σ^0_{can} is the total backscatter, σ^0_{veg} is the vegetation contribution; σ^0_{soil} is the soil contribution; τ^2 is a two-way attenuation factor; θ is the incidence angle; V_1 and V_2 are vegetation descriptors; and A and B are empirical parameters that depend on the vegetation descriptor and the radar configuration considered. It must be noted that in this study the WCM was fitted to VV polarized Sentinel-1 backscatter observations (in linear units) after terrain flattening, so γ^0 was used instead of σ^0.

In the literature, different vegetation descriptors for the WCM can be found, although V_1 and V_2 are usually considered the same (V_1=V_2). The most frequently used vegetation descriptors are: the vegetation water content (VWC) (Ayari et al., 2021), the leaf area index (LAI) (El Hajj et al., 2016), the normalized vegetation difference index (NDVI) (Bindlish and Barros, 2001), the normalized water difference index (NDWI) (Chai et al., 2015) and to a lesser extent SAR descriptors, such as radar vegetation index (RVI) (Li and Wang, 2018) or VH/VV ratio (Dabrowska-Zielinska et al., 2018). In this study, two optical-based WCM variants were implemented: one based on NDVI and the other on NDWI (Gao, 1996). Additionally, a third WCM variant was used based on the VH/VV. VH/VV values, in linear units, were smoothed with Savitzky-Golay filter (p=1, w=45) and used as the SAR descriptor, from here onwards referred to as VH/VV, for simplicity.

Although some authors used given vegetation parameter values (Wang et al., 2020) in WCM, most studies determined the values of parameters A and B for each particular site through optimization (Álvarez-Mozos et al., 2005) in a training scheme, where the soil component of backscatter (σ^0_{soil}) is an input. Some studies obtained σ^0_{soil} from physical backscattering models, such as the Integral Equation Model (IEM) (Baghdadi et al., 2017) or Dubois model (Xing et al., 2019). However, these models need additional field measurements, e.g. soil roughness. Alternatively, it can be assumed that σ^0_{soil} depends on volumetric soil moisture (SM) given Attema and Ulaby., (1978):

$$\sigma^0_{soil} = Dexp^{(SM \cdot C)} \quad (5.5)$$

where C and D are soil parameters depending on surface soil roughness and texture, respectively. In this study, only SM measurements were available and, therefore, equation 5 was used to obtain σ^0_{soil}. The only inputs used were: the VV backscatter time series in linear units (σ^0_{can}), the incidence angle, SM measurements and the vegetation descriptors considered on each variant. With this data, a non-linear overdetermined system with four unknown parameters (A, B, C and D) was stablished (Zhuo et al., 2019). The system was solved using the least-squared Powell's dog-leg algorithm (Powell, 1970), establishing the conditions that A and B are positive.

To train and validate the WCM a six-fold cross-validation scheme was implemented, which allowed to obtain A, B, C and D parameters for all fields, and state-of-the-art model performance metrics: RMSE, Pearson correlation (R) and BIAS (Han et al., 2020). Model train and validation was done separately for each WCM variant (NDVI, NDWI and VH/VV), assuming similar soil texture and roughness for all fields. For each case, the final validation results were the mean performance metrics of the six folds. The parameters of the models and the validation results can be found in Supplementary Materials 4.

Once the WCM parameters were obtained, σ^0_{soil} was extracted from equation (2) for each case, which represented the soil backscatter contribution corrected for the vegetation effect.

d) Correlation analysis

A correlation analysis was carried out to evaluate the performance of the vegetation correction methodologies investigated. For this, the original VV backscatter time series and the time series corrected with the proposed approach and the three WCM variants were correlated with the ground measured SM values, using Pearson correlation coefficient R.

Additionally, the correlation between short-term backscatter changes with sort-term SM changes was explored, since some SM retrieval methods are based on change detection techniques (Balenzano et al., 2011b). For this, the Pearson correlation coefficient of the backscatter difference between subsequent Sentinel-1 observations ($\Delta\gamma^0$) and their SM difference (ΔSM) was evaluated.

5.4.2. Results and discussion

a) Dynamics of the vegetation descriptors

The dynamics of NDVI and NDWI were quite similar (Figure 5.7), increasing right after sowing. They reached a saturation plateau that extended until vegetation ripening (Lawrence et al., 2014). This plateau was shorter for fields 5 and 6 because they were sown later, and thus had a shorter cycle. On the other hand, the dynamics of VH/VV time series were slightly different. VH/VV started to increase later, illustrating sensitivity to vegetative growth until mid-April approximately (Khabbazan et al., 2019; Meroni et al., 2021; Veloso et al., 2017). Then, there was a variable behavior at the end of the cycle. NDVI can be considered a proxy of biomass (Bindlish and Barros, 2001) and NDWI a proxy of VWC (Jackson et al., 2004). VH/VV ratio was also found to be a good proxy of crop biomass (Veloso et al., 2017).

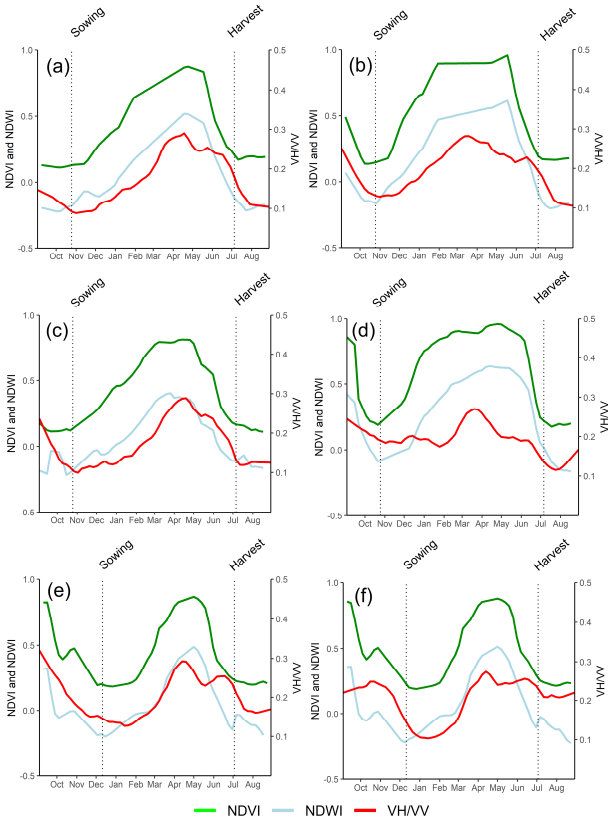

Figure 5.7. Vegetation indices time series used as vegetation descriptors in WCM. (a) Field 1, (b) Field 2, (c) Field 3, (d) Field 4, (5) Field 5and (f) Field 6. For VH/VV data from orbit 8DESC is represented, the other two have a very similar pattern

b) Corrected backscatter time series

Figure 8 shows the time series of all fields for 8DESC orbit (the results for the other two orbits are in Supplementary Materials 5). Although the amount of rainfall in spring was high, backscatter in the original time series decreased, demonstrating once again the attenuation caused by wheat in this period. The corrected time series presented differences according to the methodology applied. Focusing on WCM corrections, the first variant (NDVI) lead to significantly higher backscatter values since sowing. Similarly, the third variant (VH/VV) also resulted in higher values right after sowing. In both cases, the differences with the original time series increased with wheat development. Before harvest, the differences decreased for NDVI, but for VH/VV they remained relatively high in most fields. In the second variant (NDWI) the differences between the original backscatter time series and the corrected ones were lower. In this case, backscatter remained almost the same during the initial and last stages of the cycle, and variations only affected the attenuation period. After WCM correction with any of the three variants, it was still possible to possible to distinguish a residual attenuation in most fields (Figure 5.8).

The differences between the three WCM variants were based in the differences between the vegetation descriptors. In general, the optical corrected backscatter series started to increase too early in the season, even before the attenuation period started. Then, due to the loss of sensitivity to further vegetation growth in the plateau, the moment of maximum attenuation was not adequately corrected. Finally, in the ripening stage, vegetation indices decreased rapidly leading to a sudden drop of backscatter. This behavior was more marked in the WCM variant with NDVI than with NDWI, since the latter produced a smaller correction. The WCM based on VH/VV produced a similar result and was not able to completely correct wheat attenuation at the maximum attenuation point.

WATCOR, compared to the WCM variants investigated, only corrected backscatter values during the attenuation period, and the correction effect was smoother than that achieved with the WCM, in particular during the beginning of the season, when sudden variations in vegetation descriptors (in particular NDVI) lead to rapid changes. Furthermore, the new method adequately corrected the backscatter values at the highest attenuation point, compared to WCM corrections that were unable to completely correct it and resulted in a residual attenuation in this period.

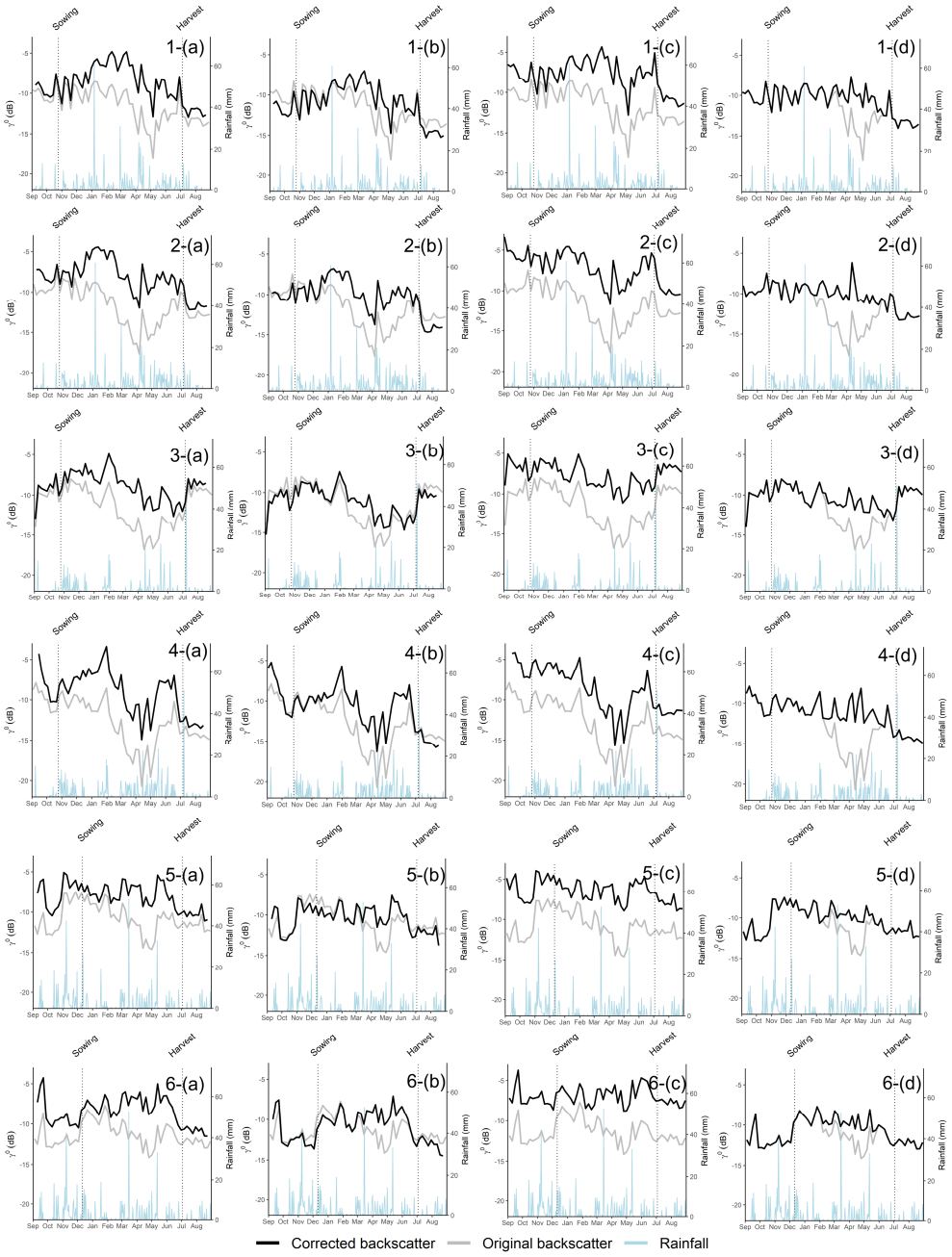

Figure 5.8. Original and corrected backscatter time series for fields 1-6 for orbit 8DESC. The backscatter corrections are: (a) WCM-NDVI, (b) WCM-NDWI, (c) WCM-VH/VV and (d) WATCOR. Rainfall includes irrigation in irrigated fields

c) Correlation results

The correlation coefficients obtained between backscatter and ground measured SM are displayed in figure 9. Additionally, scatterplots in figure 10 show the correlation obtained for all the fields together. Both figure 5.9 and 5.10 represent the correlation between γ^0 and SM and that of $\Delta\gamma^0$ and ΔSM. Further correlation results are given in Supplementary Materials 6

Regarding the correlation between γ^0 and SM, the original backscatter time series had the lowest results (Figure 5.9a), with a median R value of 0.14. In this case, individual R values obtained for the different fields and orbit passes varied strongly; with several fields even leading to negative correlations, while only a couple of fields achieved R values above 0.5. The WCM corrections based on optical data (NDVI and NDWI) improved the original correlation values, with a median value ~0.30, being the NDVI correction slightly superior to the NDWI one.

The individual results (Supplementary Materials 6) showed evidence that in most cases NDVI and NDWI had similar correlations. Other studies found that variants based on NDWI achieved better SM estimations than the ones on NDVI (Bao et al., 2018; Wang et al., 2020), attributing this result to the higher sensitivity of the SWIR channel to the vegetation water content, and the low sensitivity of NDVI to further vegetation growth after an NDVI of 0.8 (Baghdadi et al., 2017). Conversely, Zhang et al. (Zhang et al., 2021) obtained limited results with NDWI in dense wheat parcels with a modified version of the WCM, attributing this result to an eventual saturation to further increases in vegetation water content.

The correction based on the VH/VV ratio yielded lower R values, with a median value of 0.24, being only superior to the optical corrections in a few fields. The use of VH/VV as a vegetation descriptor in the WCM is appealing, as it could provide a means for correction with no need of external (optical) data. However, not many studies attempted this, and results were rather poor, with better results being achieved with NDVI at L-Band over wheat (Wang et al., 2021) and wetlands (Dabrowska-Zielinska et al., 2018).

WATCOR achieved the highest correlations, with a median value R=0.47 (Figure 5.9a). The method succeeded at increasing the correlation in most fields-orbits, with a couple of exceptions where it gave similar results to the original time series. Comparing WATCOR with the second best correction (NDVI), its performance was better in most cases. One of the exceptions was field 1, where correlation values were already high regardless of the correction applied, so differences between WCM and WATCOR were small. It must be admitted, that in some fields, in particular the two fields monitored in

2019, correlation results were low in all cases, suggesting that some external factors (e.g., soil stoniness or a later sowing) might be playing a role.

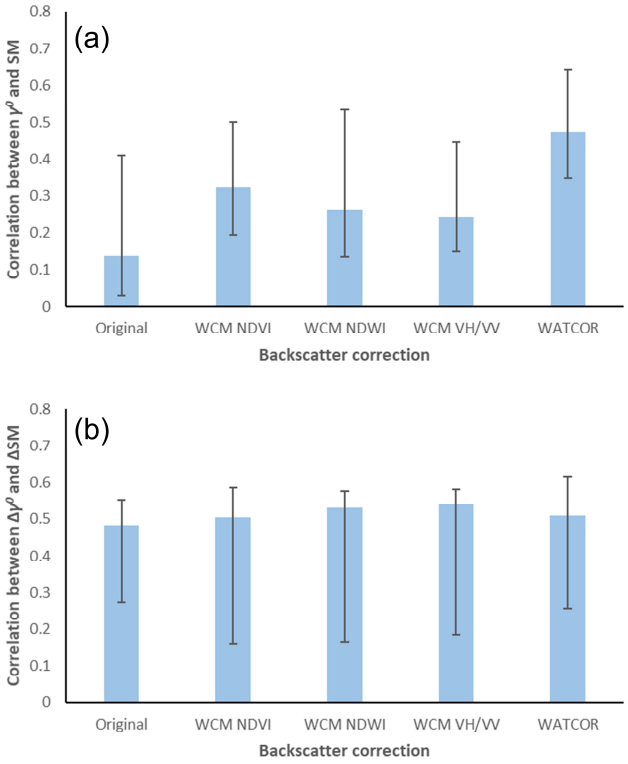

Figure 5.9. Median correlation results, (a) correlation between backscatter and SM, (b) correlation between backscatter differences ($\Delta\gamma^0$) and SM differences (ΔSM) between consecutive days. Error bars represent the first and third quartile.

When comparing the results obtained with the different orbits (Supplementary Materials 6), it can be seen that 103ASC produced poorer result in 2019, which might be due to the longer revisit time (due to the unavailability of Sentinel-1A in this orbit for that particular year), and hence a reduced sample for model fitting and correlation evaluation. In contrast, WATCOR achieved higher correlation values for 103ASC in 2018 (revisit time 6 to 12 days); and 2020 (6 days). Image acquisition in descending orbits was at early morning hours where dew was often present, which could affect backscatter (Wood et al., 2002) and its correlation with SM. WATCOR specially succeed at improving the correlation for irrigated fields compared with rain-fed ones, although it is not easy to grasp why.

Exploring the scatterplots of Figure 5.10(a-e), it can be observed that the original time series (a) had many low backscatter values (<-15dB) regardless of the SM value of that moment. All the vegetation corrections applied were successful at reducing this effect. Overall, the 0-SM correlation values always increased when a vegetation correction was applied, which confirms the necessity for correcting backscatter observations before attempting to retrieve SM over wheat fields.

Concerning the $\Delta\gamma^0$ and ΔSM correlation, it appeared that all the explored casuistry achieved similar general results. The scatterplots in figure 5.10(f-j) showed a similar behavior and correlation values in figure 5.9b (~0.51) showed no improvements after vegetation correction. Therefore, these results suggest that vegetation correction might not be necessary when applying SM retrieval methods based on change detection techniques (e.g., (Wang et al., 2020)), at least if the temporal resolution is ~6 days.

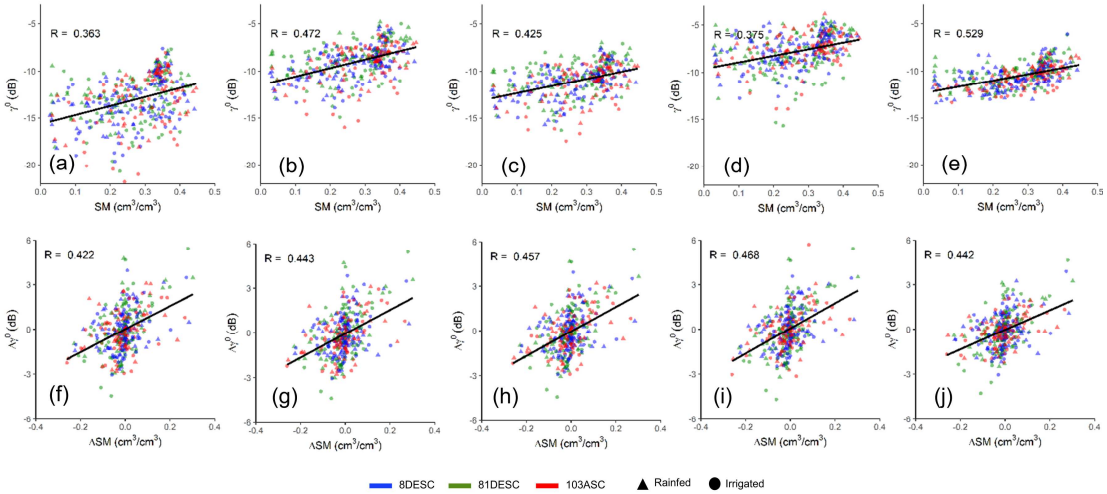

Figure 5.10. Dependence of backscatter on volumetric soil moisture for all sub-fields: absolute values (upper row) and differences between consecutive values (lower row); for the different backscatter correction investigated (a, f) original backscatter, (b, g) WCM-NDVI, (c, h) WCM-NDWI, (d, i) WCM-VH/VV (d, j) WATCOR

These results indicate that WATCOR can be useful for correcting wheat attenuation at C-band VV backscatter time series (e.g., Sentinel-1), before SM is retrieved with methods designed for bare soil conditions (Bauer-Marschallinger et al., 2019; Fung et al., 1992). The method proposed is based on sound time series analysis techniques like time series smoothing and changing point identification. Time series smoothing has been previously applied for a plethora of remote sensing applications that rely on multitemporal observations, such as crop phenology monitoring (Canisius et al., 2018; Schlund and Erasmi, 2020), reconstruction of missing data (Chen et al., 2004), land-cover classification (Shao et al., 2016) or for obtaining high quality vegetation

descriptors for climate modelling (Yuan et al., 2011). The results obtained might change if the parameters of the S-G filter were different. In this study, these parameters were selected through visual interpretation. Yet, in future studies, it would be interesting to devise an objective criterion for their selection, e.g. through cal/val schemes using additional ground truth SM datasets.

WATCOR assumes that the effect of wheat in VV backscatter is a smooth, low frequency trend, and its application requires little user intervention. It is simple, does not require external data and the results obtained even overpassed the WCM. Furthermore, unlike the WCM, WATCOR does not require any model parameterization. The only input required is the specification of approximate beginning and end dates of the attenuation period, and therefore, a previous knowledge of the study area regarding the phenological development of wheat. Therefore, the approach can be easily implemented in automatized pipelines, as a preprocessing step before SM estimation is attempted However, WATCOR relies on the whole backscatter time series, and in its present form cannot be applied in real-time. Further research is needed to improve this.

The underlying idea of WATCOR, i.e., to extract a low-frequency trend that is later subtracted from the original backscatter time series, could be transferred to other types of crops. However, this would require specific adaptations depending on the particular backscatter behavior of each crop. Furthermore, additional research is recommended to test the applicability of WATCOR at other microwave bands, as the signal penetration into the vegetation canopy is different.

5.5. Conclusion

In this paper, the backscatter behavior of wheat for C-band VV polarization was analyzed as a pre-requisite for the development of a methodology for correcting the attenuation from vegetation. A large dataset containing EU CAP declarations from farmers of four agricultural campaigns in a region (Navarre) with a high diversity of agro-climatic conditions was used. The performance of the new correction was evaluated with an analysis of the correlation between backscatter and volumetric soil moisture. The correlation was compared to WCM corrections based on optical and radar vegetation descriptors.

The results revealed that wheat time series had a typical seasonal trend, independent from soil moisture variations, with wheat plants mainly attenuating VV polarized backscatter from phenological stage BBCH21 until harvest. The behavior was identical for different agro-climatic regions, with only some differences in the amplitude of attenuation. Using standard techniques of time series smoothing and

changing point detection, it was possible to identify the start and end of the attenuation period, where the proposed method, WATCOR, was applied.

The corrected time series showed that WATCOR could effectively remove the attenuation pattern, while the time series obtained with WCM still presented a certain degree of attenuation. After correcting the vegetation effect, correlations between backscatter and soil moisture improved in all cases, with WATCOR obtaining the best general performance. Looking at results in detail, it appeared that the temporal resolution of the time series affected the results, with poorer results being observed for ASC103 pass in 2019, where revisit time was 12 days.

Overall, WATCOR provided good results and might be an alternative to other vegetation correction methods for wheat. The method does not account for the direct backscatter contribution of vegetation, contrary to the WCM, but the results suggest that this effect might be neglected in wheat, at least at C-band and VV polarization. The main advantage of the method proposed is that it does not require external information or any model parametrization. It only needs an approximate prior knowledge on the period where attenuation is likely to occur in a specific region or site. However, it relies on the whole time series of radar backscatter, making it difficult to correct the attenuation in real-time In the near future, the application of this method for soil moisture estimation techniques should be evaluated.

CHAPTER 6

EVALUATION OF SOIL MOISTURE ESTIMATION TECHNIQUES BASED ON SENTINEL-1 OBSERVATIONS OVER WHEAT FIELDS

Abstract

Soil moisture (SM) is a key variable in agriculture and its monitoring is essential. SM determines the amount of water available to plant, having a direct impact on the development of crops, on the forecasting of crop yields and on the surveillance of food security. Microwave remote sensing offers a great potential for estimating SM because it is sensitive to the dielectrical characteristics of observed surface that depend on surface soil moisture. The objective of this study is the evaluation of four different methodologies for SM estimation over wheat at field scale based on Sentinel-1 time series. Three change detection methods were selected: Short Term Change Detection (STCD), TU Wien Change Detection (TUWCD) and Multitemporal Bayesian Change Detection (MTBCD), and a Machine Learning technique: Support Vector Regression (SVR). Different methodological issues were evaluated for the implementation of these techniques at field scale. Soil moisture measurements from eight experimental wheat fields were used for the validation of the methodologies. The use of canopy backscatter observation or vegetation corrected backscatter values was also evaluated. In addition, the influence of different factors such as the satellite pass, the type of management, the month of the year and the SM content was also analyzed. The results were rather variable, with some experimental fields achieving successful performance metrics and some others rather poor ones. In general, it was observed that both TUWCD and MTBCD methods obtained better results when run with vegetation corrected backscatter time series. On the other hand, STCD and SVR produced similar results with and without vegetation correction. Both the month and the soil moisture content had an influence on the accuracy of the different methodologies. In most cases soil moisture was overestimated for dry conditions, and, to a lesser extent, it was underestimated for wet conditions.

Keywords: soil moisture, wheat, Sentinel-1, change detection, field scale

6.1. Introduction

Soil moisture (SM) is a key variable for understanding, modeling and forecasting different processes occurring at the Earth surface (Brocca et al., 2018; Green et al., 2019; Liu et al., 2020; Seneviratne et al., 2010; Wasko and Nathan, 2019). Due to its relevance, in 2010 it was recognized as an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS) of the United Nations Framework Convention on Climate Change (UNFCCC). Its monitoring is therefore necessary to track our changing climate and design proper mitigation and adaptation measures. In agricultural systems, SM determines the amount of water available to plants, and therefore, it has a direct impact on the development of crops, on the forecasting of crop yields and more importantly on the surveillance of food security (Burke and Lobell, 2010).

Remote sensing has a great potential for SM retrieval due to its capability to observe large areas of the territory repeatedly over time. Although optical and thermal sensors have potential for SM estimation (Verstraeten et al., 2006; Wang and Qu, 2007), most progress in the last decades has been achieved by the microwave scientific community (Entekhabi et al., 2015; Kerr et al., 2001; Wagner et al., 1999), due to the long-time known sensitivity of microwave reflectance and emittance to the dielectrical characteristics of observed surfaces, which mostly depend on surface SM (Ulaby et al., (1982-86.).

In recent years, global SM products at coarse spatial resolution (25-50 km) have been developed, principally based on radiometers or scatterometers (Brocca et al., 2011; Chan et al., 2016; Kerr et al., 2012; Naeimi et al., 2009). Synthetic aperture radar (SAR) sensors achieve finer spatial resolutions (10-20 m) that might be suitable to work at the scale of agricultural fields. This possibility is very appealing as it would enable a field scale SM monitoring, with direct implications for irrigation management, harvest forecast and disease control. However, SM estimation at this scale is still challenging (Peng et al., 2021), because SAR sensors are also sensitive to other terrain variables related to vegetation (Bindlish and Barros, 2001) or soil surface roughness (Verhoest et al., 2008). Therefore, the inversion of classic bare soil backscatter models, such as the physically based Integral Equation Model (IEM) (Fung, 1994) or the semi-empirical models of Oh (Oh et al., 1992), Dubois (Dubois et al., 1995) or Shi (Shi et al., 1997), is generally ill-posed and its operational application not guaranteed. Moreover, SM estimation in vegetated conditions requires the coupling of backscatter models for bare soil and for vegetation (Zhang et al., 2021). One of the most popular models is the semi-empirical Water Cloud Model (WCM) (Attema and Ulaby, 1978), which requires external vegetation descriptors and a specific parameterization for the local conditions.

This case is even more complex, although recent progress is being made in the transferability of model parameters to the regional scale (Benninga et al., 2022).

The launch of the Sentinel-1 satellites in 2014 and 2016 opened new possibilities for the estimation of SM at high spatial resolution. The unprecedent compromise between high spatial resolution, frequent revisit time and radiometric accuracy (Torres et al., 2012), along with the open data distribution policy and the operational vocation of the Copernicus program, fostered the development of new methods for SM estimation, or the adaptation of existing ones, to the characteristics of Sentinel-1 data. In particular, the development of retrieval methods based on change detection techniques that evaluate backscatter changes between consecutive observations has made significant progress. The main hypothesis of these methods is that if time series are dense enough, the backscatter differences between consecutive observations might only be caused by SM variations, since the other variables affecting backscatter, e.g. soil roughness and vegetation, could be considered constant during such a short period. Following this idea, different approaches have been developed, such as the Short Term Change Detection (STCD) approach (Balenzano et al., 2011, Balenzano et al., 2021) or the TU Wien Change Detection (TUWCD) model (Wagner et al., 1999, Bauer-Marschallinger et al., 2019). On the other hand, some other methods following different principles have also been applied to Sentinel-1 data, such as Bayesian approaches (Notarnicola et al., 2006) or machine learning algorithms like random forest regression (Liu et al., 2021), support vector regression (SVR) (Pasolli et al., 2011) or artificial neural network (ANN) (Baghdadi et al., 2012).

At present, some of these approaches are applied to routinely produce SM products at scales of ~1 km (Balenzano et al. 2021; Bauer-Marschallinger et al., 2019), which significantly improve the spatial resolution of radiometer or scatterometer based products (Zappa et al., 2022). However, for some agricultural applications such as irrigation scheduling, this spatial resolution might still be too coarse, and SM values at the field scale are ideally sought (Gao et al., 2018; Modanesi et al., 2022; Le Page et al., 2020). Furthermore, field size varies significantly around the World (White and Roy, 2015) and many agricultural areas particularly in Africa, Asia and Europe are smallholders (< 2 ha) (Lesiv et al., 2019), where high resolution is a necessity for remote sensing to be useful at all.

Therefore, the applicability and the performance of different SM estimation approaches at the field scale still need to be evaluated. The objective of this work is to evaluate different SM estimation approaches based on Sentinel-1 data for wheat fields. Three change detection approaches were selected: the Short Term Change Detection (STCD) approach (Balenzano et al., 2011a), the TU Wien Change Detection (TUWCD) model Wagner et al. (1999) and a multitemporal Bayesian change detection (MBCD)

algorithm (Notarnicola, 2014). Additionally, a machine learning approach, SVR, was also evaluated because it does not require a large training dataset (Pasolli et al., 2011), contrary to other machine learning algorithms. The approaches are calibrated on eight experimental fields with available SM measurements, and some methodological adaptations are proposed to enhance their applicability to the particular case of wheat fields. Furthermore, the influence of vegetation on backscatter, and hence on SM retrievals, was accounted for by comparing the results obtained with the original Sentinel-1 time series and those obtained after applying a vegetation correction method (WATCOR, Arias et al. (2022)) recently proposed by the authors.

6.2. Materials and methods

6.2.1. Study area and soil moisture measurements

The study focused in eight winter wheat test fields located in the province of Navarre (Spain). (Fig 6.1). Navarre is a small but diverse province regarding climate, topography and land use. Therefore, it is divided in seven agricultural regions (Arias et al., 2020), where conditions for crop growth are expected to be rather constant. Two test fields were located in region R3 (rather wet: average annual temperature of 12.5ºC and annual rainfall of 900 mm), four in region R5 (intermediate: average annual temperature of 12.7ºC and average annual rainfall of 655 mm) and two in region R6 (dry: average annual temperature of 14ºC and average annual rainfall of 350 mm). Two fields were monitored per agricultural year (2017-2018: fields 1 and 2; 2018-2019: fields 3 and 4; 2019/2020: fields 5 and 6; 2020/2021: fields 7 and 8). Each year, field pairs were managed exactly the same in terms of soil preparation, sowing and agricultural management (fertilization and other agrichemicals), except for irrigation, with half of them (odd field numbers) rainfed and the other half irrigated (even field numbers). Table 6.1 provides additional details.

Several capacitance SM probes (Sentek Sensor Technologies, Stepney SA 5069, Australia) were installed on each field in winter and they were removed just before harvest (Table 6.1). These probes recorded volumetric SM every 30 minutes at 6 different depths, from the soil surface to 60 cm deep. The top-most (0-10 cm) measurements were used in this study. For each field, the median SM time series of all the probes installed on it was calculated. Then, these time series were confronted with Sentinel-1 acquisition dates and times, selecting only the measurements closest to each Sentinel-1 acquisition. The final number of measurements used, depended on the availability of Sentinel-1A and Sentinel-1B on each of the three orbits that over-flew the study area, i.e. 103ASC, 008DESC and 081DESC, at best ~30 acquisitions were available per orbit (Table 6.2).

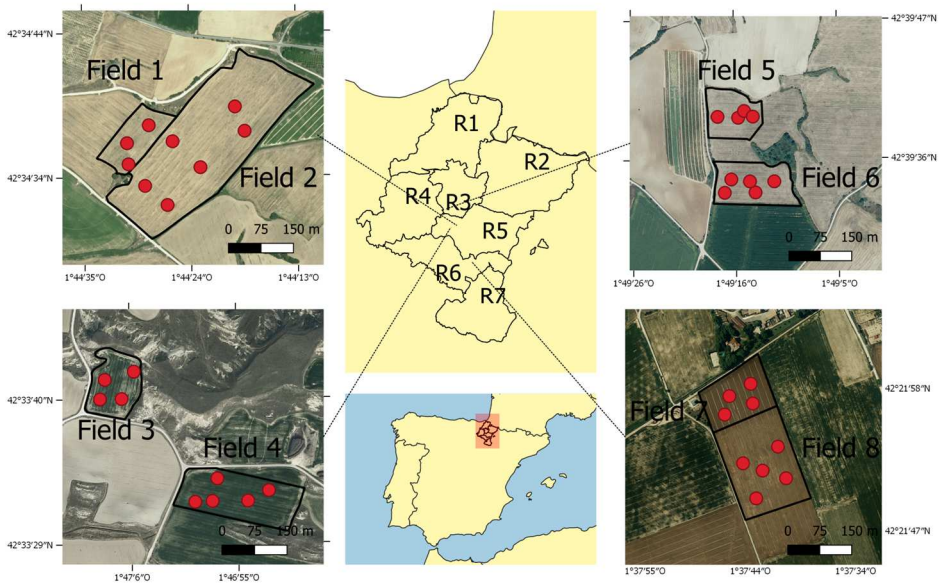

Figure 6.1. Location of the test fields and agricultural regions of Navarre. Dots represent the location of soil moisture probes.

Table 6.1. Characteristics of the test fields. Rainfall and irrigation data comprises the period sowing-harvest.

ID	Area (ha)	Texture class	% Sand	% Clay	Type	Rainfall + irrigation (mm)	Sowing	Harvest	Number of probes
1	1.36	Clay-loam	21.2	39.5	Rainfed	537	25/10/2017	05/07/2018	3
2	8.5	Loam	31.2	26.7	Irrigated	637	25/10/2017	05/07/2018	6
3	1.57	Clay-loam	23.4	31.7	Rainfed	331	25/10/2018	06/07/2019	4
4	2.8	Silt-loam	20.6	26.7	Irrigated	515	25/10/2018	06/07/2019	5
5	1.37	Clay	27.9	53.3	Rainfed	352	11/12/2019	03/07/2020	4
6	1.86	Loam	30.8	26.9	Irrigated	438	11/12/2019	03/07/2020	5
7	1.85	Silty clay loam	10.8	28.2	Rainfed	307	04/11/2020	15/07/2021	4
8	3.64	Silty clay loam	9.3	34.3	Irrigated	607	04/11/2020	15/07/2021	5

Table 6.2. Number SM measurements coincident with Sentinel-1 acquisitions per orbit; and start and end dates of SM measurements.

ID	8DESC	81DESC	103ASC	Initial date	Final Date
1-2	20	21	13	03/03/2018	27/06/2018
3-4	19	20	9	14/02/2019	21/06/2019
5-6	29	29	29	31/12/2019	28/06/2020
7-8	30	31	32	27/12/2020	02/07/2021

Surface SM dynamics varied quite strongly in the different test fields (study years) (Figure 6.2). Rain-fed fields normally experienced a transition from wet conditions in winter to dry soils in May-June (e.g., fields 1 and 7). However, this was not always the case, and field 3 had already quite low moisture conditions in winter. Field 5, in turn recorded frequent and significant precipitations during the spring, so no clear drying was observed. In general, irrigation management in fields 2, 4, 6 and 8 avoided soil drying, however, keeping in mind that measurements were taken at the soil surface, quite rapid dynamics were observed, with SM increasing rapidly due to irrigation events and decreasing also quite rapidly afterwards.

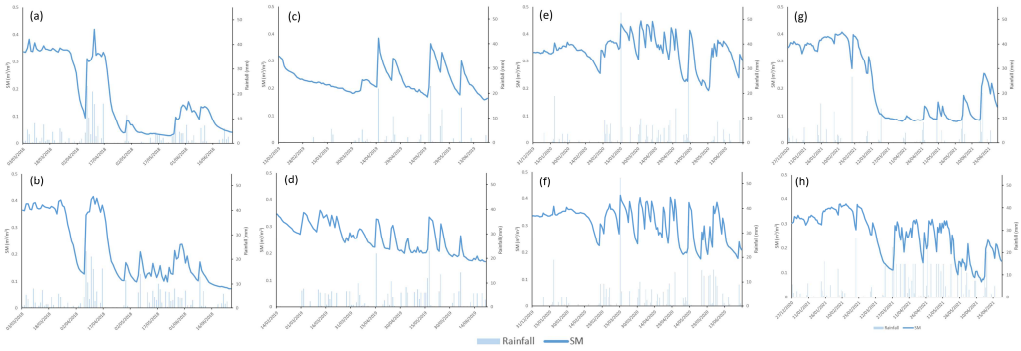

Figure 6.2. Daily median volumetric SM at 10 cm depth (lines) and daily rainfall and irrigation (bars) for the eight test fields: (a) Field 1, (b) Field 2, (c) Field 3, (d) Field 4, (e) Field 5, (f) Field 6, (g) Field 7, (h) Field 8. Fields in the top row were rainfed and those in the bottom row irrigated.

6.2.2. Satellite imagery and data extraction

Figure 6.3 summarizes all the satellite imagery and data needed for implementing the SM estimation techniques investigated, detailing the general and specific characteristics of each methodology. In the following sections, further details and explanations are given.

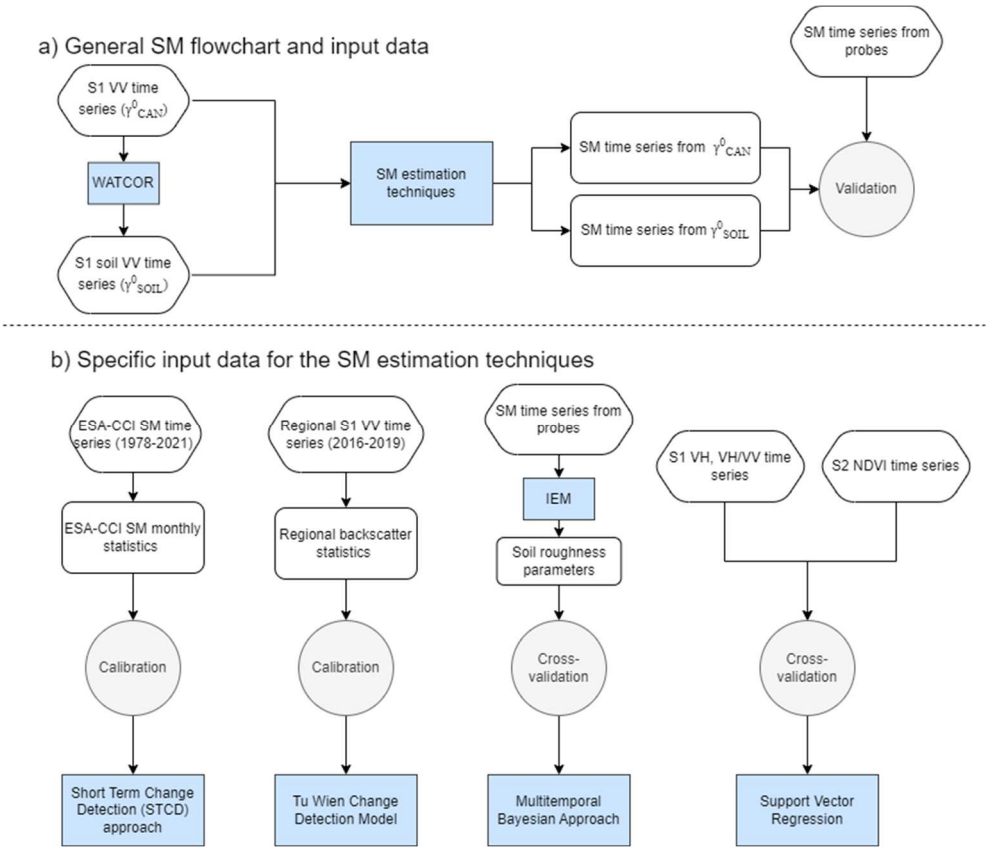

Figure 6.3. a) General SM flowchart and input data and b) specific input data for the SM estimation techniques

a) Sentinel-1

Sentinel-1 C-band SAR data was the base of this study, in particular the Interferometric Wide (IW) swath mode data with dual-pol (VH-VV) configuration, which is the pre-defined observation scenario over land. All available scenes from 1/September/2015 to 31/August/2021 covering the province of Navarre were downloaded as level-1 Ground Range Detected (GRD) products. The images corresponded to one ascending orbit (103ASC) and two descending orbits (8DESC and 81DESC).

Images were processed with an automated pipeline implemented in SNAP Graph Processing Toolbox, which followed these steps: 1) thermal noise removal; 2) slice assembly; 3) apply orbit file; 4) calibration; 5) speckle filtering (3x3 Gamma-Map); 6) terrain flattening; 7) range-doppler terrain correction and 8) subset to the extent of

Navarre. The terrain flattening and terrain correction step employed the SRTM 1sec HGT DEM. As a result, γ^0 backscatter coefficient images in VH and VV polarizations and dB units were produced, with a pixel size of 20 m. The projected local incidence angle was also obtained, as a secondary output for each scene.

For the eight test fields median backscatter coefficient values were extracted for each Sentinel-1 acquisition, resulting in backscatter time series in VV and VH polarizations. Prior to this, a 10 m inner buffer was applied to field boundaries to avoid mixed pixels. Two further processes were applied to the backscatter time series, firstly, the local incidence angle was normalized to a reference angle of θ=40^0 by applying an adapted version (Arias et al., 2022a) of the methodology proposed by Mladenova et al. (2013). Secondly, the influence of wheat canopy in VV backscatter was corrected for with the Wheat Attenuation Correction (WATCOR) method (Arias et al., 2022b), obtaining a new VV backscatter time series corresponding to the soil, named γ^0_{SOIL}, which is assumed to be free of the influence of wheat vegetation cover, as opposed to the original backscatter time series γ^0_{CAN}, which also includes the influence of wheat canopy.

b) Sentinel-2

Sentinel-2 multispectral imagery was used to obtain time series of the Normalized Vegetation Difference Index (NDVI) (Rouse et al., 1974), which was used as ancillary data for the Support Vector Regression (see section 1.3.4). In this case, the image catalogue available in Google Earth Engine (GEE) (Gorelick et al., 2017) was used, in particular, the level-2A surface reflectance collection produced by ESA. Images from 1/September/2017 to 31/August/2021 were used, discarding any acquisitions where the test fields (Table 1) were covered by clouds, shadow or snow. The median NDVI time series of each field was also obtained using the buffered polygon vector files.

c) ESA-CCI soil moisture product

The European Space Agency (ESA) provides a coarse resolution SM product through its Climate Change Initiative (CCI) (Dorigo et al., 2017). This ESA-CCI SM product is a global product that provides daily SM estimates from 1978 until 2021 at a spatial resolution of 0.25°, based on active (scatterometers) and passive (radiometers) microwave sensors. In this study, the v06.1 level 3 combined SM product was used as ancillary data for the Short Term Change Detection (STCD) approach. With this aim, the ESA-CCI SM time series of the pixels covering the test fields were downloaded and processed to obtain monthly statistics, in particular, the minimum, maximum and percentiles P05, P15, P25, P75, P85 and P95.

d) Regional characterization of wheat backscatter

For calibrating the TU Wien Change Detection Model (see section 2.3.2), a regional characterization of wheat backscatter was carried out. With this aim, a GIS database containing the field boundaries of all the EU Common Agricultural Policy (CAP) declarations in Navarre for years 2016, 2017, 2018 and 2019 was used, which was provided by the agricultural Department of the Government of Navarre. Wheat fields were extracted from this database and processed as follows: 1) create subsets for agricultural regions R3, R5 and R6; 2) mask out fields smaller than 0.5 ha; 3) apply a 5 m inner buffer to the field boundaries; 4) compute the median VV backscatter time series per field; 5) remove outliers (i.e., the 10% of fields most dissimilar to the median time series of all the fields in a region) (Arias et al. 2020); 6) normalize backscatter to a local incidence angle of θ=400 and apply WATCOR. After this process, backscatter time series of 13,200 wheat fields were obtained for region R3, 18,994 for region R5 and 11,401 for region R6; with this data a regional characterization of wheat backscatter was obtained by computing different percentiles of backscatter time series for each region (R3, R5 and R6) and orbit pass (103ASC, 8DESC and 81DESC).

6.2.3. Soil moisture estimation techniques

a) Short Term Change Detection (STCD)

The Short Term Change Detection (STCD) approach (Balenzano et al. (2011), is a change detection approach that uses dense time series (6-12 days revisit) of T co-polarized backscatter observations (γ_1^0, γ_2^0,..., γ_T^0) to obtain a SM time series. The hypothesis of the methodology is that SM changes occur at a shorter temporal scale (days) than other parameters affecting the backscatter response, (e.g. soil roughness, vegetation biomass or canopy structure), which vary at a longer temporal scale (weeks). Therefore, the ratio between two subsequent SAR observations (γ_2^0 /γ_1^0) in linear units can be expressed as a function of the dielectric constant ε and the local incidence angle θ on each observation date (Eq. 6.1).

$$\frac{\gamma_2^0}{\gamma_1^0} \approx \left|\frac{\alpha_{pp,2}(\varepsilon,\theta)}{\alpha_{pp,1}(\varepsilon,\theta)}\right|^2 \qquad (6.1)$$

Where α_{pp} is the Fresnel reflection coefficient at HH or VV polarization, and α_{VV} is defined as follows:

$$|\alpha_{VV}(\varepsilon,\theta)| = \left|\frac{(\varepsilon-1)(sin^2\theta-\varepsilon(1+sin^2\theta))}{(\varepsilon\, cos\theta+\sqrt{\epsilon-sin^2\theta})^2}\right| \qquad (6.2)$$

For N observations there are N – 1 equations with N unknown Fresnel coefficients. The system can be solved by applying a bounded least-squares optimization (Balenzano et al., 2013), where the values of the boundary conditions (α_{min}, α_{max}) play an important role in the retrieval accuracy (He et al., 2017). Once that the α values are determined for the observations, the values of ε can be derived and converted into volumetric SM values using the empirical expression of Hallikainen et al. (1985).

Palmisano et al. (2018) suggested sub-dividing the complete backscatter time series T in smaller blocks to avoid to error propagation into SM estimations. Therefore, a sliding window of N backscatter observations was considered, and the average value of the estimations was calculated (Shi et al., 2021). Different values of N (4, 5, 6, 8, 12 and 18) were evaluated and the optimum was selected.

Ideally, the boundary conditions (α_{min}, α_{max}) should correspond to the dynamic range of SM for the study area during the period of observation (Ouellette et al., 2017). Some studies constrained the problem considering fixed SM values (Balenzano et al., 2011b), others used SM field measurements (Zhang et al., 2018), and others coarser scale SM values obtained from scatterometers or radiometers (Al-Khaldi et al., 2019; Ouellette et al., 2017). In this study, the latter approach was followed, and thus the (α_{min}, α_{max}) boundary conditions were derived from the ESA CCI SM product (see section 2.2.3). With this aim, the monthly minimum and maximum ESA CCI SM values were initially considered (Supplementary materials 1.1.), but they showed a rather low dynamic range, so additional schemes were implemented enhancing the ESA CCI SM dynamic range. For this, the mean monthly difference between the maximum and minimum SM was considered, and its half was named as SM_{diff} (Eq. 6.3). This value was added or subtracted to the different ESA CCI SM monthly percentiles (Table 6.3).

$$SM_{diff} = \frac{\overline{(SM_{max.month} - SM_{min,month})}}{2} \qquad (6.3)$$

Table 6.3. Calibration schemes based on α constraints derived from ESA CCI SM

Scheme	Min SM value	Max SM value
A	Min	Max
B	P25 - SM_{diff}	P75 + SM_{diff}
C	P15 - SM_{diff}	P85 + SM_{diff}
D	P05 - SM_{diff}	P95 + SM_{diff}

The calibration of this approach consisted in evaluating which combination of schemes (Table 4) and N value provided the best results. Furthermore, all these combinations were applied to the original backscatter time series (γ^0_{CAN}) and to the one obtained after applying WATCOR (γ^0_{SOIL}). Scheme D and N=4 was chosen as the best option with almost the same results for γ^0_{SOIL} and γ^0_{CAN} (Supplementary materials 2.1.)

b) TU Wien Change Detection Model (TUWCD)

The TU Wien Change Detection Model (TUWCD) (Wagner et al. (1999) interprets changes in backscatter in a time series as changes in soil moisture, while other surface properties (geometry, roughness, vegetation, etc.) are considered as static. This model is used to produce a global SM product at a resolution of 1 km (Bauer-Marschallinger et al., 2019). Here the algorithm is applied at the field scale, using Sentinel-1 backscatter time series of each field (see section 2.2.1). This model defines the surface soil moisture content (SSM) as a relative index between 0 and 1, which is estimated from the backscatter value of that particular day normalized with some minimum and maximum backscatter boundary conditions (Eq. 6.4).

$$SSM(t) = \frac{\gamma^0(t) - \gamma^0_{min}}{\gamma^0_{max} - \gamma^0_{min}} \qquad (6.4)$$

where $\gamma^0(t)$ is the backscatter observation in dB units at time t, and γ^0_{min} and γ^0_{max} are the minimum and maximum backscatter values in dB units corresponding to dry and saturated soil conditions. These values need to be extracted from long time series where it is likely that the pixel or polygon of interest would reach these dry and saturated conditions some time (Wagner et al., 1999).

Bauer-Marschallinger et al. (2019) already mentioned that the relatively short length of the Sentinel-1 data record might result in an absence of the dry and saturated conditions required to successfully apply this algorithm. This might be particularly difficult in humid regions (Zribi et al., 2014) or in agricultural areas with irrigation systems, where completely dry conditions might never be met. Furthermore, in agricultural areas, more and more often managed under crop rotation schemes (European Comission, 2021), yearly changing vegetation covers might also influence backscatter dynamics (Veloso et al., 2017; Arias et al., 2020) making the selection of γ^0_{min} and γ^0_{max} very challenging. To overcome this problem, this study proposes a regional characterization of backscatter for wheat fields (Section 2.2.4). For each agricultural region, backscatter time series of thousands of wheat fields were obtained for four different years and their statistics (minimum, maximum and different percentiles) were computed and used to calibrate the algorithm, that is, to select the dry and wet references, evaluating different schemes (Table 6.4).

Table 6.4. Schemes for obtaining the dry and wet soil conditions (γ^0_{min}, γ^0_{max}) in the TUWCD model from the regional statistics of wheat fields time series.

Scheme	γ^0_{min}	γ^0_{max}	Scheme	γ^0_{min}	γ^0_{max}
1	min	max	9	P0.01	max
2	min	P0.999	10	P0.01	P0.999
3	min	P0.99	11	P0.01	P0.99
4	min	P0.9	12	P0.01	P0.9
5	P0.001	max	13	P0.1	max
6	P0.001	P0.999	14	P0.1	P0.999
7	P0.001	P0.99	15	P0.1	P0.99
8	P0.001	P0.9	16	P0.1	P0.9

These different calibration schemes were applied to both γ^0_{CAN} and γ^0_{SOIL}, to investigate the eventual benefits of correcting for the vegetation influence with WATCOR. Backscatter values below or above the limits were set to 0 and 1 respectively (Hornáček et al., 2012). Then, relative SSM values were linearly scaled to SM values (m^3/m^3) (Carranza et al., 2019):

$$SM(t) = \left(SM_{sat} - SM_{wp}\right) * SSM(t) + SM_{wp} \qquad (6.5)$$

where SM_{sat} is the saturated soil moisture content and SM_{wp} is the wilting point, both estimated from soil texture data (Saxton and Rawls (2006); Rawls et al. (1982)). The results of this calibration (see Supplementary materials 2.2.) showed that the optimum results were obtained for scheme 3.

c) Multitemporal Bayesian Change Detection approach (MTBCD)

The multitemporal Bayesian Change Detection approach (MTBCD) is an inversion procedure for SM estimation based on the Bayes' theorem, adapted from Notarnicola, (2014). The objective is to infer the unknown soil dielectric constant time series $(\varepsilon_1, \varepsilon_1, \dots, \varepsilon_T)$ from the available Sentinel-1 backscatter time series $(\gamma^0_1, \gamma^0_2, \dots, \gamma^0_T)$. By applying Bayes' theorem, it is possible to turn probabilities estimated from a training dataset into probabilities for the estimation of the unknown variable ε (Gelman et al., 1995).

The conditional probability density function (pdf) $P(\gamma^0_1, \gamma^0_2, \dots, |\varepsilon_i)$, which is the probability of finding the vector of γ^0_i given specific values of ε_i, is estimated from a training set of backscatter observations and their corresponding ground measurements of ε. By using the Integral Equation Model (IEM) (Fung, 1994), the theoretical backscatter values $\gamma^0_{i,IEM}$ calculated from the ground SM measurements are obtained. These values are compared to the Sentinel-1 backscatter values introducing a random

variable R. This variable accounts for the sensor noise and model errors (Notarnicola et al., 2008):

$$\gamma_i^0 = R\gamma_{i,IEM}^0 \qquad (6.6)$$

The pdf of *R* is assumed to follow a Gaussian distribution (Eq. 7), and its mean (μ) and standard deviation (σ), are determined by using the maximum likelihood principle. Tests were carried out to check whether this distribution represents this kind of data (Notarnicola et al., 2008).

$$P(R) = \frac{e^{-(R-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \qquad (6.7)$$

Once that the pdf parameters are calculated, the Bayes' theorem from two consecutive γ_1^0, γ_2^0 observations for obtaining the conditional density function $P(\varepsilon|\gamma_1^0, \gamma_2^0)$ is applied:

$$P(\varepsilon|\gamma_1^0, \gamma_2^0) = \frac{P_{prior}(\varepsilon)P_{post}(\gamma_1^0, \gamma_2^0|\varepsilon)}{P(\gamma_1^0, \gamma_2^0)} \qquad (6.8)$$

where:

- P_{prior} is the a priori joint density function for ε. A uniform density function over the physical range of the parameter can be assumed.
- $P_{post}(\gamma_1^0, \gamma_2^0|\varepsilon)$ is the posterior density function based on measured values.
- $P(\gamma_1^0, \gamma_2^0)$ is a normalization factor.

This function can be expressed in terms of the probability density $P(R)$ by a transformation detailed in Notarnicola et al. (2006). The optimal estimator $\bar{\varepsilon}$ for ε, that has de minimum variance is the conditional mean:

$$\bar{\varepsilon} = \frac{\int(\varepsilon-\varepsilon_{min})P_{prior}(\varepsilon)\left(\frac{1}{\gamma_{IEM}^0}\right)^2 P(\frac{\gamma_1^0}{\gamma_{IEM}^0})P(\frac{\gamma_2^0}{\gamma_{IEM}^0})d\varepsilon}{P(\gamma_1^0, \gamma_2^0)} \qquad (6.9)$$

where ε_{min} is the minimum value of the uniform density function from P_{prior}.

This approach needs to be trained and validated. The IEM requires as input the values of ε and the soil roughness parameters *s* (standard deviation of heights) and *l* (correlation length). The unavailability of soil roughness measurements in this study required the optimization of *s* and *l*. Due to the fact that all the backscatter values were terrain flattened and normalized to a common incidence angle, it was assumed that the incidence angle was the same for all fields. In total, 414 combinations of soil roughness parameters were evaluated:

- *s*: from 0.2 cm to 2 cm with a step of 0.1 cm
- *l*: from 2 cm to 25 cm with a step of 1 cm.

For each combination, the backscatter values were simulated using as input the ground SM measurements.

The optimization function selected the combination of roughness parameters that minimized the root mean-squared error (RMSE) between the observed backscatter values, and the simulated γ^0_{IEM} values. The lowest RMSE was achieved with *s*=0.4 and *l*=3 for γ^0_{SOIL}. (Supplementary Materials 2.3.).

A 4 fold-cross validation scheme was performed separately for each orbit pass. In each fold, 6 fields were used in the training phase and 2 fields were used for validation. Therefore, the SM was estimated in all fields. The training phase consisted in determining the pdf parameters that were used later for SM estimation in the validation fields.

d) Support Vector Regression (SVR)

Vapnik (1995) proposed a machine learning technique called Support Vector Regression (SVR) that has become very popular for empirically predicting different variables (Cervantes et al., 2020), and among others SM (Ahmad et al., 2010; Pasolli et al., 2011). This algorithm is able to handle complex and nonlinear problems (Pasolli et al., 2015b) even if the size of the dataset is small (Bruzzone and Melgani, 2005). SVR transforms nonlinear problems into linear problems in the high-dimensional space, using a kernel function that replaces the inner product operation in the high-dimensional space (Chen et al., 2021). The types of kernel functions that can be adopted are polynomial, linear, sigmoid and Gaussian. Gaussians radial basis functions (RBF) are usually used due to their good performance and simplicity (Bruzzone and Melgani, 2005). SVR results depend on the parameterization of three parameters: the RBF kernel parameter (Gamma), the regularization parameter (Cost) and the width of the insensitive tube (Epsilon).

SM estimation with SVR consists of a first training part where the best parameters have to be selected for the model construction, and a validation part with an independent test dataset. In this study, the scheme for training and validation was the same as the one used in the MTBCD approach. The SVR was implemented in R using the library “e1071”.

SVR training was performed separately for each set of input variables (Table 6.5). The numerical input variables were normalized between 0 and 1 and the categorical variable “month” was transformed to binary vectors using One-Hot encoding. The grid

search strategy with a 5-fold cross validation was used for the optimization of the model parameters, based on the Mean Squared Error (MSE). The ranges for the parameters were [10^{-3},10^{3}], [10^{-3},10^{3}] and [10^{-3},10] for Gamma, Cost and Epsilon, respectively. Then, the optimal parameters were used for training the model and the SM was estimated with the validation dataset.

Table 6.5. SVR schemes based on different input variables.

Scheme	Input variables	Scheme	Input variables
1	γ^0_{CAN}	12	γ^0_{SOIL}, NDVI
2	γ^0_{SOIL}	13	γ^0_{CAN}, γ^0_{VH}, VH/VV, NDVI
3	γ^0_{VH}	14	γ^0_{SOIL}, γ^0_{VH}, VH/VV, NDVI
4	VH/VV	15	γ^0_{CAN}, γ^0_{SOIL}, γ^0_{VH}, VH/VV, NDVI
5	γ^0_{CAN}, γ^0_{VH}	16	γ^0_{CAN}, month
6	γ^0_{CAN}, VH/VV	17	γ^0_{SOIL}, month
7	γ^0_{SOIL}, γ^0_{VH}	18	γ^0_{CAN}, NDVI, month
8	γ^0_{SOIL}, VH/VV	19	γ^0_{CAN}, γ^0_{VH}, VH/VV, month
9	γ^0_{CAN}, γ^0_{VH}, VH/VV	20	γ^0_{SOIL}, γ^0_{VH}, VH/VV, month
10	γ^0_{SOIL}, γ^0_{VH}, VH/VV	21	γ^0_{CAN}, γ^0_{SOIL}, γ^0_{VH}, VH/VV, month
11	γ^0_{CAN}, NDVI	22	γ^0_{CAN}, γ^0_{VH}, VH/VV, NDVI, month

For further analysis, the best schemes for γ^0_{CAN} and γ^0_{SOIL} based solely on Sentinel-1 were selected (Supplementary materials 2.4). Therefore, the outputs of scheme 19 (γ^0_{CAN}, γ^0_{VH}, VH/VV and month) and scheme 20 (γ^0_{SOIL}, γ^0_{VH}, VH/VV and month) were chosen.

6.2.4. Evaluation of results

The performance of the evaluated SM retrieval techniques was assessed with different metrics calculated between the volumetric SM recorded by the probes (SMobs) and the estimated SM (SMest) (Entekhabi et al., 2010): Pearson Correlation (R), root-mean-square error (RMSE), BIAS and unbiased root-mean-square error (ubRMSE). The metrics were calculated separately for each test field and Sentinel-1 orbit.

6.3. Results

In this section, the statistical results for the two schemes selected for each methodology (Section 2.3 and Supplementary Materials) are presented. The numerical results for every field and orbit pass are found in Supplementary Materials 3.

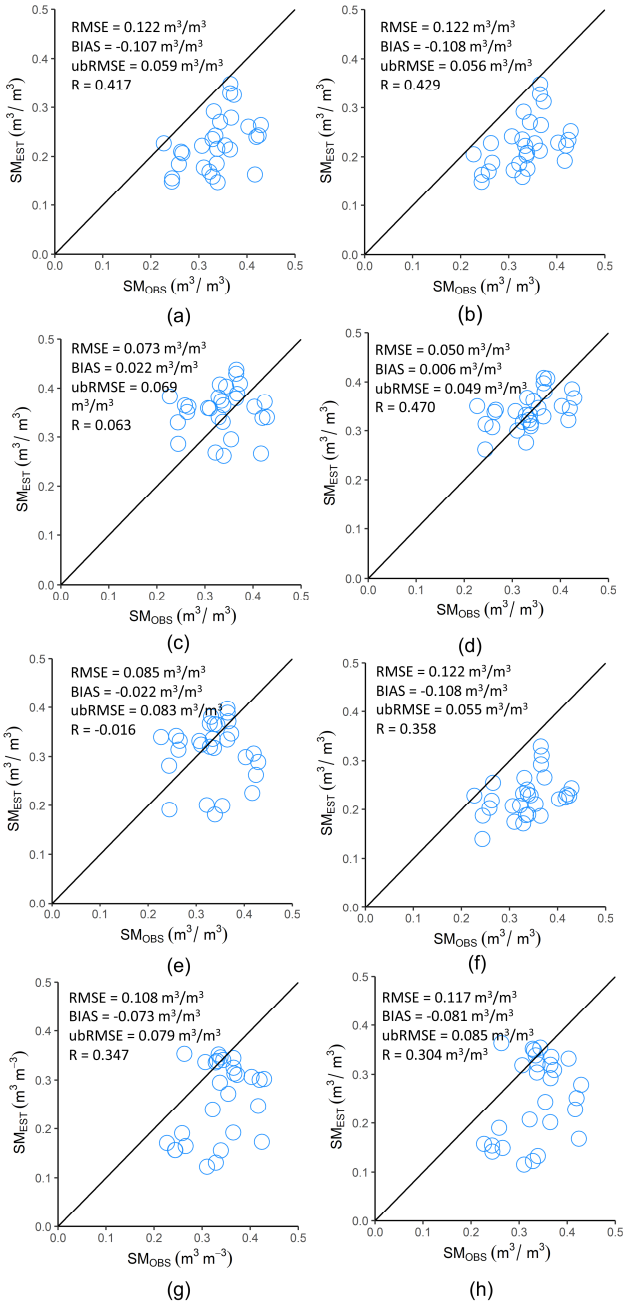

Figure 6.4. Scatterplots of estimated SM as a function of in situ SM measurements for the different methodologies for a sample plot (field 5 and orbit pass 8DESC). (a) STCD (γ^0_{CAN}); (b) STCD (γ^0_{SOIL}); c) TUWCD (γ^0_{CAN}); (d) TUWCD (γ^0_{SOIL}); (e) MTBCD (γ^0_{CAN}); (f) MTBCD (γ^0_{SOIL}); g) SVR (γ^0_{CAN}); (h) SVR (γ^0_{SOIL}).

6.3.1. Comparison between methodologies

In figure 6.4, the scatterplots between observed and estimated SM after applying the four approaches to one of the test fields and one descending pass is represented, using as input both the γ^0_{CAN} and γ^0_{SOIL} time series. In this case, the best results are obtained with TUWCD and γ^0_{SOIL}. For the rest of the fields, the performance metrics are summarized in the boxplot of Figure 5.

Figure 6.5 shows the results after computing the statistical values per orbit pass and field. Points not only represent the different test fields (colors), but also the Sentinel-1 orbit pass used in each case (symbols). Boxplots summarize all the point data, showing the median and quartiles of the field results for each approach.

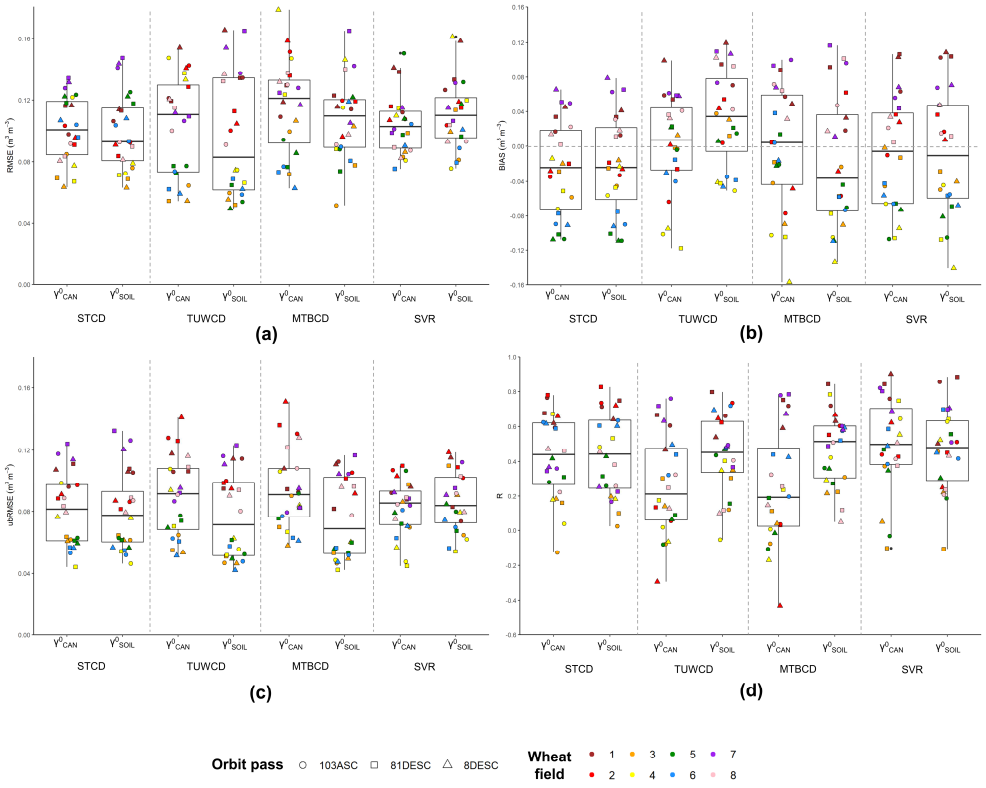

Figure 6.5. Statistical results for the different SM estimation techniques. (a) RMSE; (b) BIAS; (c) ubRMSE; (d) Correlation.

Median values of the performance metrics demonstrate a positive effect of the vegetation correction in all techniques except for the SVR. In particular, γ^0_{SOIL} metrics improved those obtained with γ^0_{CAN} in the TUWCD and MTBCD approaches and to a lesser extent in the STCD. Different metrics provide different views on the performance of the techniques, regarding the RMSE best results were obtained with the TUWCD (0.08 m^3/m^3), followed by STCD (0.09 m^3/m^3), SVR (0.10 m^3/m^3) and MTBCD (0.12 m^3/m^3). However, the SVR had the lowest bias, whereas STCD and MTBCD had negative bias and TUWCD a positive one. Looking at the unbiased ubRMSE, the TUWCD approach also provided the best results (0.06 m^3/m^3) but practically matched by the MTBCD, and then followed by STCD and SVR (both with 0.08 m^3/m^3). In turn, highest correlation values were obtained by the MTBCD approach (0.52), followed by SVR (0.49), TUWCD (0.46) and STCD (0.45).

Looking at field results in detail, it can be observed that results varied strongly from cases to cases, with rather large variability ranges for all metrics: RMSE (0.05 – 0.17 m^3/m^3), BIAS (-0.16 – 0.12 m^3/m^3), ubRMSE (0.04 – 0.15 m^3/m^3) and R (-0.4 – 0.9). The approaches with the highest variability were TUWCD and MTBCD. On the other hand, SVR and STCD provided more consistent results and had a lower variability. This field-to-field variability might be related to characteristics of the fields, for instance, fields 1, 7 and 8, all had a positive BIAS in all techniques, indicating a systematic overestimation of SM. Conversely, fields 3, 4, 5 and 6 had a negative BIAS (except for TUWCD). In terms of correlation, the case of field 3 was particular, since all the methodologies provided poor performance in this field.

For a more in-depth analysis, each method with its best backscatter input option (γ^0_{SOIL} for STCD, TUWCD and MTBCD, and γ^0_{CAN} for SVR) was examined to evaluate the influence of different factors on the results. These factors were the agricultural management (rainfed or irrigated fields), the satellite orbit passes, the month of the year and the soil moisture content (low, medium, high).

6.3.2. Comparison between agricultural management

The results based on the agricultural management are shown in figure 6.6. RMSE metrics did not differ greatly in terms of management for MTBCD and SVR (~0.10 m^3/m^3), although the variability was larger in rainfed fields. For STCD and TUWCD approaches, the RMSE was better in irrigated fields. However, the higher value of RMSE in rainfed TUWCD was caused by the high values achieved by fields 1 and 7, as fields 3 and 5 had the lowest RMSE values of all cases. Regarding the ubRMSE, STCD, TUWCD and MTBCD achieved similar median results (~0.08 m^3/m^3) regardless of the agricultural management. Specifically, fields 1, 2, 7 and 8 had higher ubRMSE (>0.08 m^3/m^3) than fields 2, 4, 6, 8 (ubRMSE ~0.05 m^3/m^3). SVR results differed slightly,

because rainfed fields presented very low variability in comparison to irrigated fields, although the median value was only slightly higher. In terms of correlation, the values were similar among methods and management practices (R=42), with the exception of the STCD where irrigated fields had higher values compared to rainfed fields.

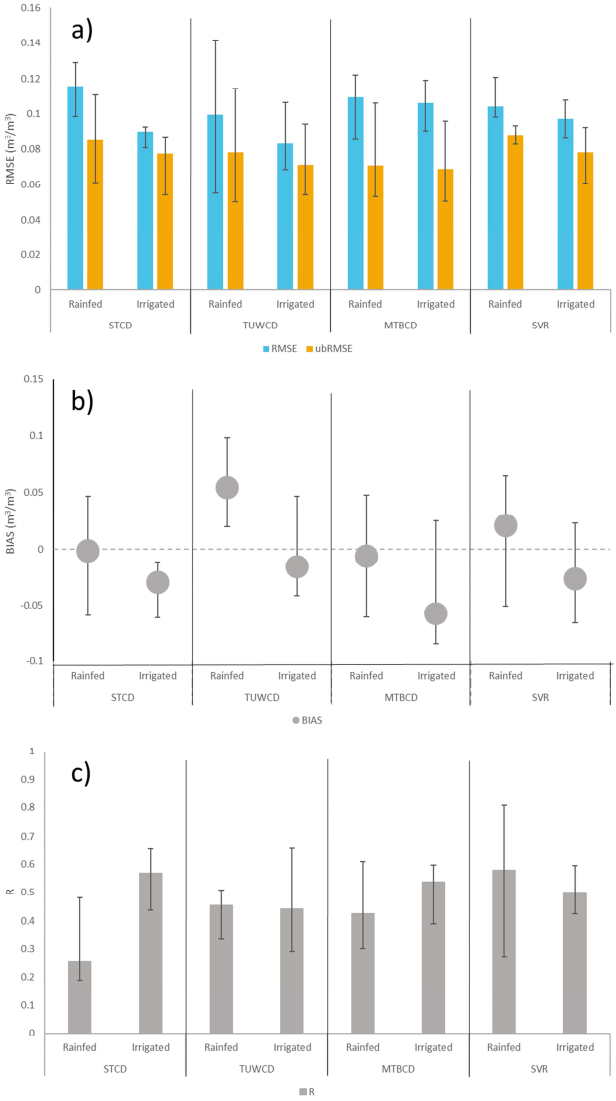

Figure 6.6. Statistical results for the different SM estimation techniques and agricultural management. (a) RMSE and ubRMSE; (b) BIAS; (c) Correlation

6.3.3. Comparison between orbit passes

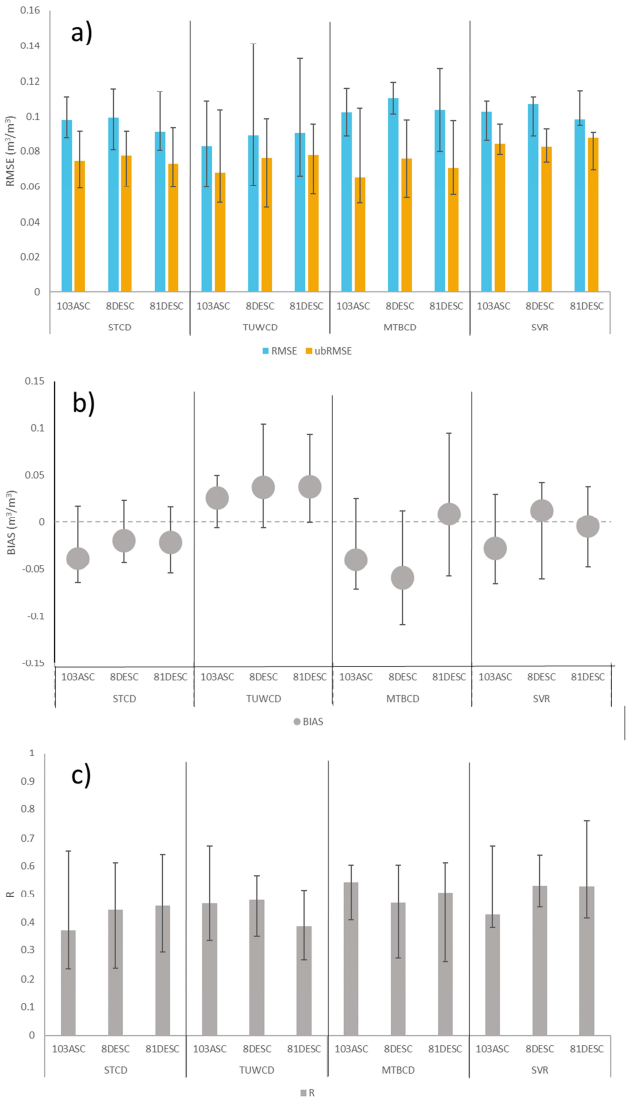

Figure 6.7. Statistical results for the different SM estimation techniques and orbit passes. (a) RMSE and ubRMSE; (b) BIAS; (c) Correlation.

The performance metrics obtained for the three orbit passes (figure 6.7) did not show clear differences. In general, RMSE values for all methods and orbit passes were similar, with a median around 0.08-0.09 m^3/m^3. The descending orbits in TUWCD and

the descending orbit 81 in MTBCD presented a larger RMSE variability that was reduced after removing the BIAS. Thus, the ubRMSE results between the orbits were very similar for all methods. Correlation results did not show either any clear influence of satellite orbit passes. For TUWCD correlation values were slightly lower for orbit 81DESC, and for MTBCD 103ASC resulted in slightly higher correlations. In any case these differences were not significant, and performance metrics seemed to vary more strongly depending on the particular field than on the orbit pass.

6.3.4. Evaluation of results by month

To evaluate results per month, the performance metrics were computed per orbit pass, field and month, and the median metrics were computed per month. Correlation results in this case are not reliable, since a reduced number of (SM_{obs}, SM_{est}) pairs were available for each aggrupation, and only RMSE results are shown (figure 6.8).

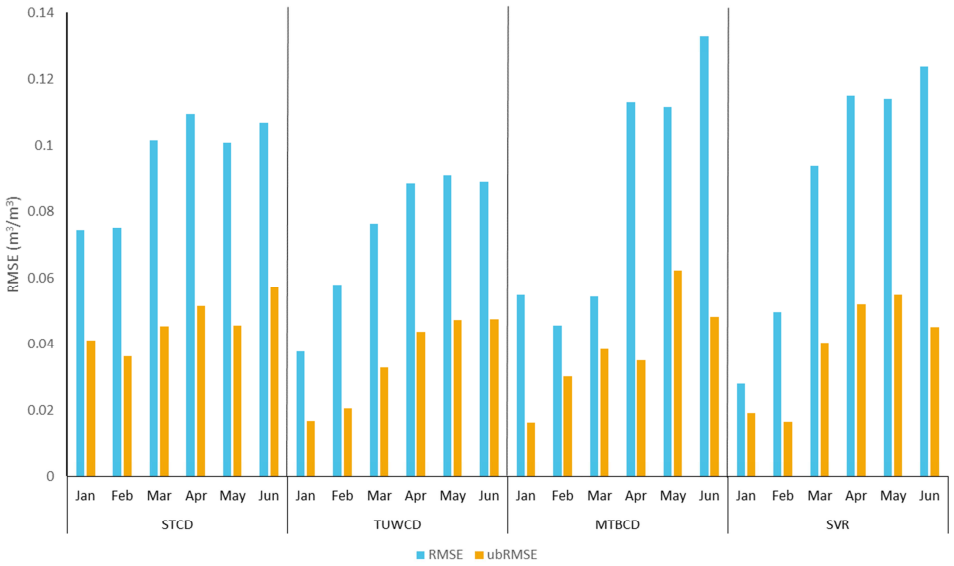

Figure 6.8. Median RMSE and ubRMSE results for the different SM estimation techniques and months.

There were clear differences depending on the month, particularly for MTBCD and SVR. RMSE values were lower in winter months (January and February) for all cases, with values lower than 0.05 m3/m3 in the TUWCD, MTBCD and SVR. In STCD, the RMSE was slightly higher (0.075 m3/m3). The RMSE in March was higher for all methods except MTBCD, and April, May and June obtained the highest values in all cases (0.08 – 0.12 m3/m3). RMSE was exceptionally high in June in the MTBCD (0.13 m3/m3). Regarding ubRMSE results, TUWCD and SVR presented very low values in January and

February (< 0.02 m3/m3), while the other months had higher values that did not exceed 0.05 m3/m3. MTBCD also had a very good result in January (< 0.02 m3/m3), mid-values in February, March and April (~0.035 m3/m3), and higher results in May and June (0.05 – 0.06 m3/m3). Finally, STCD was the approach that had overall higher ubRMSE results, but the differences between months were minor, with values ranging 0.035 – 0.055 m3/m3

6.3.5. Evaluation of results based on soil moisture content

With the aim of evaluating the performance of the retrieval methods depending on the actual soil moisture content of fields, the statistics were computed per orbit pass, field, and observed SM divided in three levels:

- Dry: SM < 0.15 m^3/m^3
- Intermediate: 0.15 m^3/m^3 <= SM < 0.30 m^3/m^3
- Wet: SM >= 0.30 m^3/m^3

The median RMSE results (figure 6.9) showed certainly high error values for dry conditions in all the techniques (0.013 – 0.18 m^3/m^3). Wet conditions also presented relatively high RMSE metrics (0.09 – 0.011 m^3/m^3) for all approaches except TUWCD that had a better performance (RMSE = 0.06 m^3/m^3). The intermediate moisture produced the best results, with values ~0.06 m^3/m^3 for the STCD, TUWCD and MTBCD, and RMSE = 0.08 m^3/m^3 for the SVR. All approaches presented a high positive BIAS (>0.12 m^3/m^3) for dry conditions, conversely BIAS was negative for wet conditions (-0.05 – -0.1 m^3/m^3). Intermediate conditions obtained the lowest BIAS for all approaches (~ -0.025 m^3/m^3, except TUWCD with 0.04 m^3/m^3). Unbiased RMSE values (ubRMSE) were similar for the four techniques (0.03 – 0.05 m^3/m^3) and did not depend on the moisture conditions.

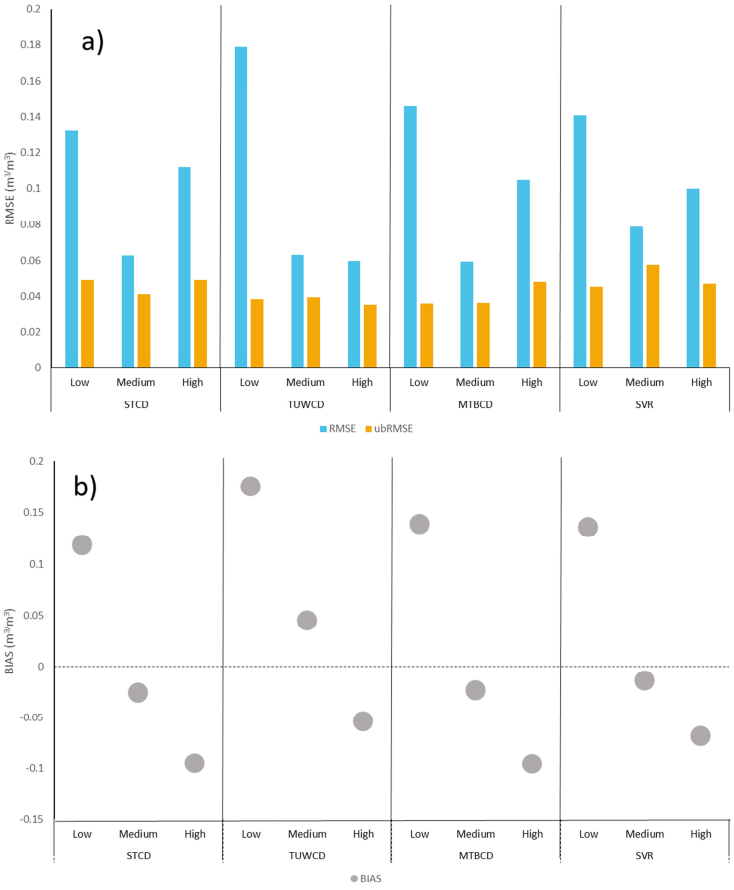

Figure 6.9. a) RMSE, ubRMSE and b) BIAS results for the different SM estimation techniques depending on the SM conditions. Bars represent median values.

These results showed that, depending on the actual SM content, the approaches could overestimate or underestimate SM. In figure 6.10, SM time series for different test fields (1, 2, 5 and 6) are displayed to see the behavior of the different approaches. Figure 2 shows that fields 1 and 2 had a similar SM content during January and February, and then, SM dropped dramatically in field 1, whereas SM remained slightly higher in field 2 due to irrigation. In this case, although the evaluated approaches were sensitive to the precipitation event and subsequent soil wetting in April 2018, they did not predict adequately the significant drying of field 1. MTBC and STCD approaches estimated SM better in the irrigated Field 2. On the other hand, fields 3, 4, 5 and 6 had rather high SM, and although field 6 was under irrigation, it showed slightly lower SM values than field 5 because precipitations were abundant that season. For these fields, TUWCD provided the best estimations, predicting a range of SM values similar to

measurements and showing sensitivity to some wetting and drying events throughout the season. STCD and MTBC also showed some sensitivity to SM dynamics, but their SM estimates were severely biased (underestimated) SVR showed higher values during the first period of estimation in line with ground measurements, but the predictions were also very underestimated during spring months. Similarly, to field 1, field 7 also had an extreme SM drop in March, and neither of the approaches was able to estimate the low SM values from March to May. Although field 8 also presented the SM drop as field 7, the estimations were better because SM was higher during spring months due to irrigation events. In this field, TUWCD had the higher BIAS because SM level was intermediate almost the whole period.

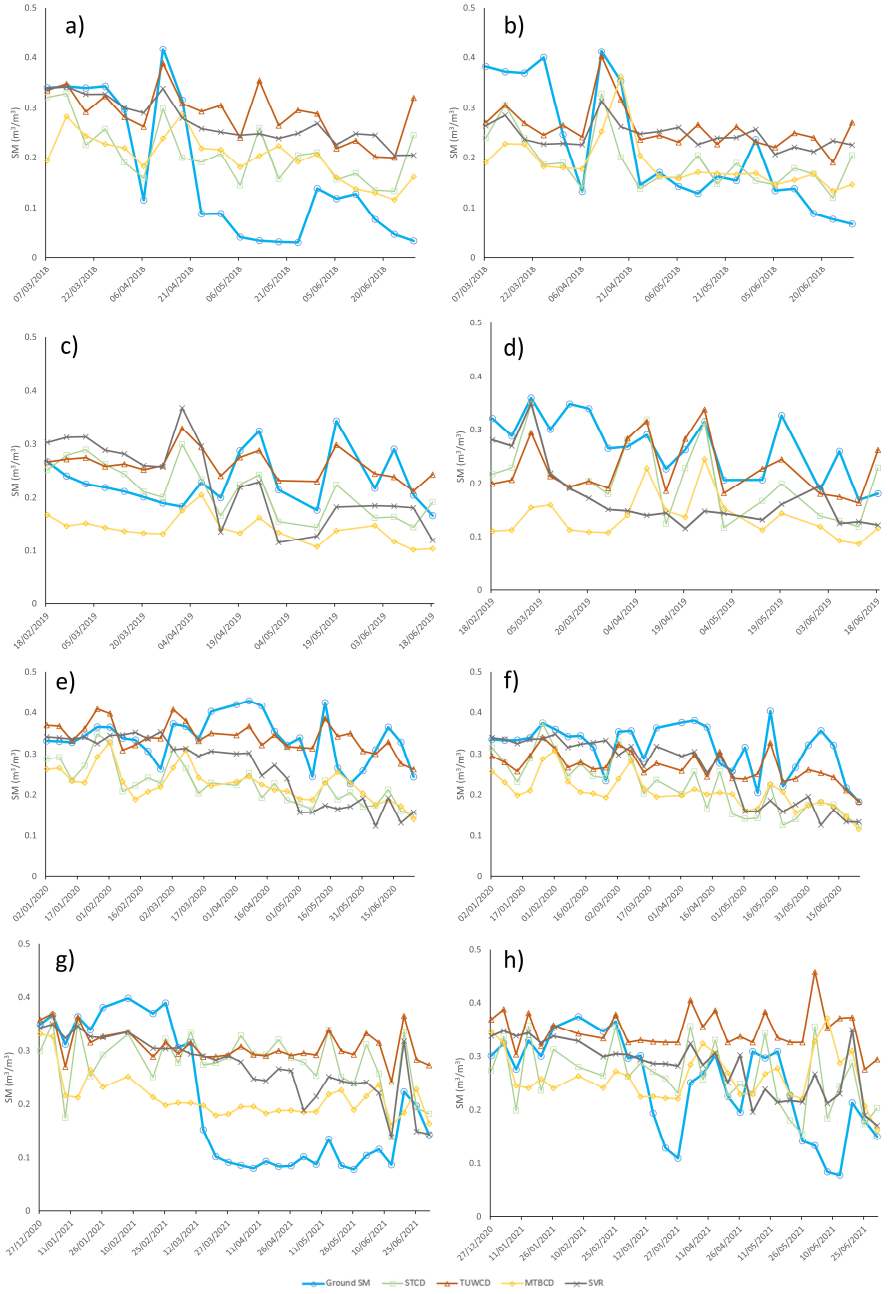

Figure 6.10. Ground SM time series and estimated SM of the different approaches with orbit 8DESC: a) Field 1; b) Field 2; c) Field 3; d) Field 4; e) Field 5; f) Field 6; g) Field 7 and h) Field 8.

6.4. Discussion

6.4.1. Calibration of methodologies

The calibration results for the STCD approach showed that, in a situation when the dynamic range of SM to be estimated is unknown, the selection of the SM bounds is critical because the accuracy of the algorithm might decrease significantly. This was also pointed out by Ouellette et al. (2017). For instance, for the best N scenario, the correlation increased from 0.3 to 0.45 for scheme A or scheme D, respectively. In terms of correlation, the size of the sliding window (N value) also had major influence on the accuracy of results, in particular for γ^0_{CAN}. As N increased, the correlation decreased, this can be interpreted as a smoothing effect of N on SM estimates, blurring the short-term SM dynamics. Other researchers stated that large N values might foster error propagation in the algorithm (Palmisano et al., 2018), and provide SM estimates affected by changing surface and vegetation conditions in the mid-term (Balenzano et al., 2011, Zhu et al., 2022). Balenzano et al. (2011) found that N = 3 achieved better results than N = 11. In our study, large N values lead to poorer results, but particularly with γ^0_{CAN}, demonstrating the mid-term influence of vegetation in backscatter and the subsequent SM retrievals. Better results were obtained with γ^0_{SOIL}.

The lack of sufficiently long and dense backscatter time series for the test fields (due in part to crop rotation) made the calibration of the TUWCD more challenging. A regional characterization of wheat fields' backscatter was implemented to obtain the required γ^0_{min} and γ^0_{max} values. This approach was highly affected by the influence of vegetation, as the correlation results for γ^0_{CAN} where ~0.20, while for γ^0_{SOIL} where ~0.45. This is explained by the backscatter behavior of wheat (Arias et al., 2022; Mattia et al., 2003). The TUWCD algorithm assumes that the minimum backscatter value corresponds to dry soil, but in wheat canopies, the minimum backscatter values are normally produced by the attenuation of the canopy at the end of stem elongation phase.

The MTBCD approach was different in terms of applicability. An optimization of the soil roughness parameters based on the IEM was performed due to the absence of field roughness measurements, and therefore, there was only a set of results for γ^0_{CAN} and γ^0_{SOIL} based on a cross-validation scheme. For this method, the results also showed that the vegetation canopy clearly affected the results, with γ^0_{SOIL} achieving better performance metrics than γ^0_{CAN}.

Different input variables schemes were investigated for the SVR approach in the calibration phase. The obtained RMSE values for the different calibration schemes did not differ much, however the correlation results change significantly. SVR models

provided best results when considering all the backscatter input variables and, very importantly, when adding the month as an input variable. Although the best correlation was achieved by also adding the NDVI, the selected schemes were based only on radar features, as NDVI might not be always available. Furthermore, the results for the best non-NDVI schemes were almost as good. Similarly, Holtgrave et al. (2018) found that after applying the SVR for SM estimation in vegetated areas with S1, the accuracy of SM retrieval decreased only 1.05% vol with only SAR data compared to the use of SAR and optical data. In this case, the vegetation correction did not provide any benefits in the estimations.

6.4.2. Comparison of methods

The comparative analysis between the four methodologies and the two variants (γ^0_{CAN} and γ^0_{SOIL}) after calibration demonstrated clear benefits after vegetation correction for TUWCD and MTBCD. In STCD and SVR, the results were similar for both backscatter inputs. The overall performance metrics showed that there was not any approach that clearly outperformed the rest.

RMSE results varied strongly from field to field, ranging from 0.05 m^3/m^3 to 0.16 m^3/m^3. These rather high values were also encountered in a SVR study (Attarzadeh et al., 2018). In terms of BIAS, the TUWCD appeared to have a different behavior, as positive BIAS was found in most test fields, contrary to the rest of techniques were negative BIAS predominated. The ubRMSE results were similar for the different approaches, with fields 3, 4, 5 and 6 obtaining good results (~0.05 m^3/m^3), and fields 1, 2, 7 and 8 showing poorer results (>0.08 m^3/m^3). These differences might be related to the overall moisture conditions of each field (figure 2), with higher ubRMSE values being observed for fields with more extreme (normally dryer) conditions during most of the season, contrary to the fields where SM did not drop below 0.15 m^3/m^3. The correlation results between approaches were also similar, being slightly higher in the MTBCD and worse in the STCD. In general, there were specific fields that achieved very good correlation values while others not. Specifically, field 3 had the poorer correlation results. These differences between fields might be caused by local conditions, such as soil texture, stoniness or organic matter content, which might affect SM dynamics and also vegetation growth as found in Ouaadi et al. (2020).

The statistical results of this study are comparable to other works. RMSE values were higher when using the SVR at the field scale (object-based approach) when compared to pixel-based with 20 m of spatial resolution (Attarzadeh et al., 2018). A strong field variability was also observed for RMSE, ubRMSE and R when applying the STCD at a scale of 500 m in agricultural areas (Foucras et al., 2020). Amazirh et al. (2018) obtained a RMSE of 0.16 m^3/m^3 when using TUWCD. Even at coarser

resolutions, medium correlation values and intrinsic RMSE of ~0.07 m^3/m^3 was obtained after applying STCD in larger areas (Balenzano et al., 2021b) and RMSE in the range 0.08-0.12 m^3/m^3 in vegetated areas (Pulvirenti et al., 2018). Using other SM estimation methods, similar results were also found in wheat (Ma et al., 2020) or in cropland areas (Benninga et al., 2022; Qiu et al., 2019).

The analysis of the performance of the methods based on the actual SM conditions showed that all approaches overestimated SM in dry conditions and underestimated it in wet conditions, explaining why medium-high SM fields achieved better RMSE results. Holtgrave et al. (2018) obtained similar results with a SVR model and Zhu et al. (2022) with the STCD. These authors interpreted the poor performance for wet conditions as a consequence of the lower sensitivity of SAR backscatter to high SM values. Overestimation was stronger with SVR and especially TUWCD when compared to STCD and MTBCD, which might be a decisive factor when choosing one approach over the other. The analysis per month showed better results in winter months, although these differences could be caused in part by the wetter SM conditions in this time of the year in the study area (figure 2). The SVR approach can incorporate this by adding a categorical variable 'month' as input, this resulted in improved results with this approach. The analysis between rainfed and irrigated fields and the use of different satellite passes did not show significant differences.

SM retrieval at field-scale is challenging due to the high complexity of the SAR signal (Bauer-Marschallinger et al., 2019), especially in vegetated areas, and none of the approaches analyzed in this study provided optimal results for the conditions investigated (wheat fields of 1-8 ha), especially in dry conditions. Yet, the use of these SM products could still be useful even with biases (Entekhabi et al., 2010), as long as the estimations reproduce reasonably the temporal dynamics of SM (Koster et al., 2009). The choice of one method or another is not straightforward, as the performance is similar for the different approaches. The use of STCD or SVR algorithms could be more advantageous as they might not need any vegetation correction. However, STCD performance depends on the selection of SM bounds, and SVR needs calibration with ground SM data. A limitation of this study might be the lack of ground SM data during the autumn-winter period. Further studies focused on this period might complement the results obtained here. In addition, it would be convenient to perform similar analyses.

6.4. Conclusions

In this study, four different SM estimation approaches based on Sentinel-1 data for wheat crop at field scale were evaluated. Three change detection techniques (STCD, TUWCD and MTBCD) and one machine learning technique (SVR) were selected. The approaches were tested in eight experimental fields with available SM measurements in Navarre (Spain). The influence of wheat canopy on backscatter and SM estimation was also assessed by implementing the WATCOR vegetation correction method (γ^0_{SOIL}) and comparing this results with the uncorrected backscatter time series (γ^0_{CAN}).

Regarding the implementation of each approach, the following ideas can be outlined:

- STCD approach provided best results when applied to rather short periods of time (four S-1 observations or ~1 month) and considering rather extreme boundary conditions. Furthermore, almost the same results were obtained for γ^0_{SOIL} and γ^0_{CAN}, and thus this approach might be applied in wheat fields with no previous vegetation correction.
- TUWCD required a regional characterization of wheat fields backscatter to obtain the backscatter bounds due to the lack of long and dense backscatter time series for the test fields. This method provided best results when applied to vegetation corrected γ^0_{SOIL} time series, with RMSE and particularly correlation values worsening when γ^0_{CAN} was used instead. Therefore, vegetation correction seems to be necessary in this case.
- MTBCD required an optimization of soil roughness parameters using the IEM, and best results were obtained when using γ^0_{SOIL} time series.
- SVR models provided best results when considering all the backscatter input variables and, very importantly, when adding the month as an input variable. In this case, the vegetation correction did not provide any benefits in the estimations.

The comparison between methodologies showed similar statistical metrics between the estimated and measured SM for the different approaches. Fields that presented low SM contents had a higher RMSE that was related to a SM overestimation effect for dry conditions (SM < 0.15 m^3/m^3). This overestimation was greater in TUWCD and SVR. To a lesser extent, all methodologies underestimated SM for wet conditions. The correlation values were similar for all the approaches and varied between fields. Overall, the performance of the evaluated approaches was comparable to similar studies, yet it might be emphasized that the implemented approaches were applied at the field scale (1-8 ha), and required only Sentinel-1 data as input.

CONCLUSIONS

This thesis is framed in the present context in which the agricultural sector has to face several challenges related to the World population growth and the impact of climate change in agricultural production systems. Unfortunately, these challenges are expected to worsen in the future, and remote sensing techniques will play an essential role in crop monitoring and contribute to ensure food security. Remote sensing is the ideal tool for large-scale crop monitoring, as it provides information over large areas of the territory on a regular basis. The research effort carried out in this doctoral thesis focused on the use of Sentinel-1 time series for the development, analysis and evaluation of different methodologies aimed at estimating agricultural variables of interest at the field scale.

In the first study, a supervised crop classification technique based on Sentinel-1 time series was proposed. This technique uses VH and VV polarization and VH/VV ratio time series obtained over a large number of fields in an area with high agro-climatic variability. The analysis of the backscatter signatures of the different crops showed that the VH and VV channels are sensitive to various phenological stages where particular structural or wetness changes occur in the canopy. On the other hand, the VH/VV ratio is sensitive to vegetative growth in general, providing little information for crop identification. Crops with greater singularities in their time series were classified better, and the combination of VH, VV and VH/VV time series provided the best classification results. Field size had a notable influence on the results, with larger fields obtaining better accuracies. Classification results also improved when stratifying the territory into agricultural regions, showing that agroclimatic diversity has a negative effect on performance. High proportion of fallow fields and very heterogeneous classes also decreased the classification accuracy. When defining crop classes in future studies, it is recommended to avoid very diverse classes, dividing them into classes with more homogenous crops. Nevertheless, the results obtained suggest that it is possible to implement this method in an operational manner in regions with frequent cloudiness within the framework of CAP monitoring, as well in studies pursuing other objectives. In the future, the results should be compared with those obtained with optical images or with machine learning techniques.

The second work addressed the influence of image acquisition geometry in the context of using time series acquired with different relative orbits. The first analysis evaluated the influence of the incidence angle in wheat time series for the backscatter coefficient (σ^0) and for the terrain-flattened backscatter coefficient (γ^0). The application

of terrain-flattening practically eliminated the influence of incidence angle in VH, but not in VV. In this polarization, although the effect of incidence angle in γ^0 was much lower than in σ^0, there was still some influence that was variable throughout wheat growth cycle. The influence was minimal when the crop canopy was at its maximum development, but stronger in periods were surface scattering was predominant. Thus, it is shown that the influence of incidence angle at VV polarization varies as a function of the phenological development of wheat. The evaluation of the backscatter differences between the three orbits in the study area showed that terrain-flattening is able to reduce the angular variations in backscatter data. Despite this, applying an incidence angle normalization technique to γ^0 can further reduce the differences, especially in winter and bare soil periods. After applying these two methods, the correlation with reNDVI improved, showing a better description of wheat growth. The anisotropy correction had a minor effect, only relevant in May. Differences between ascending and descending passes may be due in part to acquisition geometry, but also to the time of the day, with possible variations due to dew, soil moisture or frost. This issues need to be studied in more detail. Therefore, for studies whereby several orbits have to be combined, the need to correct γ^0 time series should be considered, as the results may improve in quantitative studies aimed at retrieving bio-geophysical variables.

In the third study, the backscatter dynamics of wheat in VV polarization was analyzed using a large field dataset. The results showed that wheat presents a characteristic long-term pattern independent of soil moisture variations, in which plants produce an attenuation of vertical waves from the tillering stage to plant senescence. The attenuation effect is proportional to wheat growth and has a maximum around the end of April, coinciding with the heading stage. Based on these results, a new methodology that corrects this attenuation was proposed. It was called WATCOR and to be applied it only requires the VV backscatter time series itself. WATCOR incorporates classical time series analysis techniques: smoothing techniques, changing point detections, linear interpolation and envelope fitting. The effectiveness of WATCOR was evaluated by means of a correlation analysis between backscatter and soil moisture measurements in several wheat fields. Additionally, several variants of the WCM, based on optical and radar vegetation descriptors, were applied to correct the effect of vegetation, comparing the results obtained with WATCOR. The results showed that WATCOR effectively eliminated the attenuation trend in the time series, while the time series obtained with the different variants of the WCM applied still showed some attenuation effect. The correlations with soil moisture increased in all cases when the corrections were applied, with WATCOR having the best general performance. Therefore, the new WATCOR method showed promising results in correcting the attenuation effect of wheat and may be a viable alternative to other vegetation correction methodologies. Although it does not take into account the direct

contribution of vegetation on backscatter, the results showed that this effect could be minimal in VV polarization. WATCOR does not require external information or parameterizations, which is an advantage if compared to other methods (e.g., WCM). However, it cannot be implemented in real time, as it requires the complete time series of the wheat growth cycle to be applied. Future research efforts should try to improve this method so that it can be applied in real time.

In the last study, four different SM estimation techniques using Sentinel-1 time series were evaluated. The effectiveness of WATCOR on the different approaches was also assessed by comparing the results obtained with the original time series and the WATCOR corrected ones. Three change detection methods (STCD, TUWCD and MTBCD) and a Machine Learning technique (SVR) were applied on eight experimental wheat fields in Navarre where SM measurements were available. Each methodology had specific considerations in the calibration or training phase. STCD obtained better results when considering short time periods (4 S1 observations) and rather extreme SM limits. Vegetation correction was not necessary in this method. On the other hand, TUWCD and MTBCD obtained better results with WATCOR corrected data, thus demonstrating the sensitivity of these techniques to the effect of vegetation. TUWCD requires establishing maximum and minimum backscatter bounds that, at the field scale, cannot be easily obtained from long time series due to crop rotation. To circumvent this a regional characterization of wheat backscatter was proposed to estimate maximum and minimum bounds. Finally, SVR showed no improvement when applying WATCOR, but the addition of the variable 'month' increased the accuracy of results. Comparisons between the best combinations for each method showed similar performance metrics. In all approaches, an overestimation of low moisture values was observed, especially in TUWCD and SVR. To a lesser extent, high SM values were also underestimated in all techniques. Despite the complexity of estimating SM at the field scale solely with SAR data, acceptable estimations were obtained. However, further research is needed to obtain more robust methods that enable the estimation of SM at the field scale.

To sum up, the results of this thesis show that the analysis and extraction of the information contained in SAR time series is useful for various agricultural applications, such as crop classification or soil moisture estimation. This can be especially interesting in areas with high persistence of cloud cover where using optical imagery is unfeasible. With all this, it is expected that the results of this work will contribute to a better understanding of the backscattering behavior of different crops, especially wheat.

CONCLUSIONES

Esta tesis se enmarca en el contexto actual en el cual el sector agrícola tiene que hacer frente a varios retos relacionados con el incremento de la población mundial y el impacto del cambio climático en los sistemas de producción agrícolas. Desafortunadamente, se prevé que estos retos se agraven en el futuro, por lo que las técnicas de teledetección desempeñarán un papen fundamental en la monitorización de los cultivos y contribuirán a garantizar la seguridad alimentaria. La teledetección es la herramienta ideal para el seguimiento de los cultivos a gran escala, ya que proporciona información sobre grandes áreas del territorio de forma regular. El esfuerzo de investigación realizado en esta tesis doctoral se ha centrado en el uso de las series temporales de Sentinel-1 para el desarrollo, análisis y evaluación de diferentes metodologías orientadas a la estimación de variables agrícolas de interés a escala de campo.

En el primer estudio, se propuso una técnica de clasificación supervisada de cultivos basada en series temporales Sentinel-1. Esta técnica utiliza series temporales en polarización VH y VV, y series temporales del ratio VH/VV obtenidas sobre un gran número de parcelas agrícolas en una zona con alta variabilidad agroclimática. El análisis de las firmas de retrodispersión de los diferentes cultivos mostró que los canales VH y VV son sensibles a varios estados fenológicos donde ocurren cambios particulares en la estructura o la humedad en las cubiertas agrícolas. Por otro lado, el ratio VH/VV es sensible al crecimiento vegetativo en general, brindando poca información para la identificación de cultivos. Los cultivos con mayores singularidades en sus series temporales se clasificaron mejor, y la combinación de series temporales VH, VV y VH/VV proporcionó los mejores resultados de clasificación. El tamaño de parcela tuvo una influencia notable en los resultados, ya que los campos más grandes obtuvieron mejores precisiones. Los resultados de la clasificación también mejoraron al estratificar el territorio en regiones agrícolas, mostrando que la diversidad agroclimática tiene un efecto negativo en los resultados. La alta proporción de campos en barbecho y las clases muy heterogéneas también redujeron la precisión de la clasificación. Al definir clases de cultivos en futuros estudios, se recomienda evitar clases muy diversas, dividiéndolas en clases con cultivos más homogéneos. No obstante, los resultados obtenidos sugieren que es posible implementar este método de manera operativa en regiones con nubosidad frecuente en el marco del seguimiento de la PAC, así como en estudios que persigan otros objetivos. En el futuro, los resultados deberán compararse con los obtenidos con imágenes ópticas o con técnicas de aprendizaje automático.

El segundo trabajo abordó la influencia de la geometría de adquisición de imágenes en el contexto del uso de series temporales adquiridas con diferentes órbitas relativas. El primer análisis evaluó la influencia del ángulo de incidencia en las series temporales de trigo para el coeficiente de retrodispersión (σ^0) y para el coeficiente de retrodispersión después de aplicar terrain-flattening (γ^0). La aplicación del terrain-flattening prácticamente eliminó la influencia del ángulo de incidencia en VH, pero no en VV. En esta polarización, aunque el efecto del ángulo de incidencia en γ^0 fue mucho menor que en σ^0, seguía existiendo cierta influencia que fue variable a lo largo del ciclo de crecimiento del trigo. La influencia fue mínima cuando la cubierta del cultivo se encontraba en su máximo desarrollo, pero fue más fuerte en los periodos en los que la retrodispersión superficial era predominante. Así, se demuestra que la influencia del ángulo de incidencia en la polarización VV varía en función del desarrollo fenológico del trigo. La evaluación de las diferencias de retrodispersión entre las tres órbitas en la zona de estudio mostró que el terrain-flatteining es capaz de reducir las variaciones angulares de la retrodispersión. A pesar de ello, la aplicación de una técnica de normalización del ángulo de incidencia a γ^0 puede reducir aún más las diferencias, especialmente en los períodos de invierno y de suelo desnudo. Tras aplicar estos dos métodos, la correlación con reNDVI mejoró, mostrando una mejor descripción del crecimiento del trigo. La corrección de la anisotropía tuvo un efecto menor, sólo relevante en mayo. Las diferencias entre las pasadas ascendentes y descendentes pueden deberse en parte a la geometría de adquisición, pero también a la hora del día, con posibles variaciones debidas al rocío, la humedad del suelo o las heladas. Estas cuestiones deben estudiarse con más detalle. Por lo tanto, para estudios en los que haya que combinar varias órbitas, debe considerarse la necesidad de corregir las series temporales γ^0, ya que los resultados pueden mejorar en los estudios cuantitativos destinados a obtener variables biogeofísicas.

En el tercer estudio, se analizó la dinámica de retrodispersión del trigo en polarización VV utilizando un gran dataset de parcelas. Los resultados mostraron que el trigo presenta un patrón característico a largo plazo, independiente de las variaciones de humedad del suelo, en el que las plantas producen una atenuación de las ondas verticales desde la fase de ahijado hasta la senescencia de la planta. El efecto de atenuación es proporcional al crecimiento del trigo y tiene un máximo alrededor de finales de abril, coincidiendo con la etapa de espigado. A partir de estos resultados, se propuso una nueva metodología que corrige esta atenuación. Se denominó WATCOR y para su aplicación sólo requiere la propia serie temporal de retrodispersión VV. WATCOR incorpora técnicas clásicas de análisis de series temporales: técnicas de suavizado, detección de puntos de cambio, interpolación lineal y ajuste de envolventes. La eficacia de WATCOR se evaluó mediante un análisis de correlación entre la retrodispersión y las medidas humedad del suelo en varios campos de trigo. Además,

se aplicaron diversas variantes del WCM, basadas en descriptores de vegetación ópticos y radar, para corregir el efecto de la vegetación, comparando los resultados obtenidos con WATCOR. Los resultados mostraron que WATCOR eliminó eficazmente la tendencia de atenuación en las series temporales, mientras que las series temporales obtenidas con las distintas variantes del WCM seguían mostrando cierto efecto de atenuación. Las correlaciones con la humedad del suelo aumentaron en todos los casos cuando se aplicaron las correcciones, siendo WATCOR el que obtuvo el mejor rendimiento general. Por lo tanto, el nuevo método WATCOR mostró resultados prometedores en la corrección del efecto de atenuación del trigo y puede ser una alternativa viable a otras metodologías de corrección de la vegetación. Aunque no tiene en cuenta la contribución directa de la vegetación en la retrodispersión, los resultados mostraron que este efecto podría ser mínimo en la polarización VV. WATCOR no requiere información o parametrizaciones externas, lo cual es una ventaja si se compara con otros métodos (por ejemplo, WCM). Sin embargo, no puede ser implementado en tiempo real, ya que para su aplicación requiere la serie temporal completa del ciclo de crecimiento del trigo. Los futuros esfuerzos de investigación deberían tratar de mejorar este método para que pueda aplicarse en tiempo real.

En el último estudio se evaluaron cuatro técnicas diferentes de estimación de humedad del suelo (HS) utilizando series temporales Sentinel-1. También se evaluó la eficacia de WATCOR en los diferentes enfoques, comparando los resultados obtenidos con las series temporales originales y las corregidas con WATCOR. Se aplicaron tres métodos de detección de cambios (STCD, TUWCD y MTBCD) y una técnica de aprendizaje automático (SVR) en ocho campos de trigo experimentales de Navarra en los que se disponía de mediciones de HS. Cada metodología tuvo consideraciones específicas en la fase de calibración o entrenamiento. El STCD obtuvo mejores resultados cuando se consideraron periodos de tiempo cortos (4 observaciones de S1) y límites de HS bastante extremos. La corrección de la vegetación no fue necesaria en este método. Por otro lado, TUWCD y MTBCD obtuvieron mejores resultados con datos corregidos por WATCOR, demostrando así la sensibilidad de estas técnicas al efecto de la vegetación. TUWCD requiere establecer límites máximos y mínimos de retrodispersión que, a escala de campo, no pueden obtenerse fácilmente a partir de series temporales largas debido a la rotación de cultivos. Para evitar esto, se propuso una caracterización regional de la retrodispersión del trigo para estimar los límites máximos y mínimos. Por último, el método SVR no mostró ninguna mejora al aplicar WATCOR, pero la adición de la variable "mes" aumentó la precisión de los resultados. Las comparaciones entre las mejores combinaciones de cada método mostraron métricas de rendimiento similares. En todos los enfoques se observó una sobreestimación de los valores de humedad bajos, especialmente en TUWCD y SVR. En menor medida, los valores altos de SM también se subestimaron en todas las técnicas.

A pesar de la complejidad de estimar la HS a escala de campo únicamente con datos de SAR, se obtuvieron estimaciones aceptables. Sin embargo, es necesario seguir investigando para obtener métodos más robustos que permitan estimar la HS a escala de campo.

En resumen, los resultados de esta tesis demuestran que el análisis y la extracción de la información contenida en las series temporales SAR son útiles para diversas aplicaciones agrícolas, como la clasificación de cultivos o la estimación de la humedad del suelo. Esto puede ser especialmente interesante en zonas con alta persistencia de nubosidad donde el uso de imágenes ópticas es inviable. Con todo ello, se espera que los resultados de este trabajo contribuyan a una mejor comprensión del comportamiento de la retrodispersión de diferentes cultivos, especialmente el trigo.

REFERENCES

Ahmad S, Kalra A, Stephen H. 2010. Estimating soil moisture using remote sensing data: A machine learning approach. Advances in Water Resources **33** : 69–80. DOI: 10.1016/j.advwatres.2009.10.008 [online] Available from: http://dx.doi.org/10.1016/j.advwatres.2009.10.008

Al-Khaldi MM, Johnson JT, O'Brien AJ, Balenzano A, Mattia F. 2019. Time-Series Retrieval of Soil Moisture Using CYGNSS. IEEE Transactions on Geoscience and Remote Sensing **57** : 4322–4331. DOI: 10.1109/TGRS.2018.2890646

Álvarez-Mozos J, Casalí J, González-Audícana M, Verhoest NEC. 2005. Correlation between ground measured soil moisture and RADARSAT-1 derived backscattering coefficient over an agricultural catchment of Navarre (North of Spain). Biosystems Engineering **92** : 119–133. DOI: 10.1016/j.biosystemseng.2005.06.008

Álvarez-Mozos J, Casalí J, González-Audícana M, Verhoest NEC. 2006. Assessment of the operational applicability of RADARSAT-1 data for surface soil moisture estimation. IEEE Transactions on Geoscience and Remote Sensing **44** : 913–923. DOI: 10.1109/TGRS.2005.862248

Amazirh A, Merlin O, Er-Raki S, Gao Q, Rivalland V, Malbeteau Y, Khabba S, Escorihuela MJ. 2018. Retrieving surface soil moisture at high spatio-temporal resolution from a synergy between Sentinel-1 radar and Landsat thermal data: A study case over bare soil. Remote Sensing of Environment **211** : 321–337. DOI: 10.1016/j.rse.2018.04.013 [online] Available from: https://doi.org/10.1016/j.rse.2018.04.013

Ardila JP, Tolpekin V, Bijker W. 2010. Angular backscatter variation in L-band ALOS ScanSAR images of tropical forest areas. IEEE Geoscience and Remote Sensing Letters **7** : 821–825. DOI: 10.1109/LGRS.2010.2048411

Arias M, Campo-Bescós MA, Alvarez-Mozos J. 2020. Crop Classification Based on Temporal Signatures of Sentinel-1 Observations over Navarre Province, Spain. Remote Sensing : 1–29. DOI: 10.3390/rs12020278

Arias M, Campo-Bescos MA, Arregui LM, Gonzalez-Audicana M, Alvarez-Mozos J. 2022. A New Methodology for Wheat Attenuation Correction at C-Band VV-Polarized Backscatter Time Series. IEEE Transactions on Geoscience and Remote Sensing **60** : 1–14. DOI: 10.1109/tgrs.2022.3176144

Attarzadeh R, Amini J, Notarnicola C, Greifeneder F. 2018. Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at plot scale. Remote Sensing **10** : 1–18. DOI: 10.3390/rs10081285

Attema EPW, Ulaby FT. 1978. Vegetation modeled as a water cloud. Radio Science **13** : 357–364. DOI: 10.1029/RS013i002p00357

Atzberger C. 2013. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. Remote Sensing **5** : 949–981. DOI: 10.3390/rs5020949

Ayari E, Kassouk Z, Lili-Chabaane Z, Baghdadi N, Bousbih S, Zribi M. 2021. Cereal crops soil parameters retrieval using L-band ALOS-2 and C-band sentinel-1 sensors. Remote Sensing **13** DOI: 10.3390/rs13071393

Baghdadi N, Choker M, Zribi M, El Hajj M, Paloscia S, Verhoest NEC, Lievens H, Baup F, Mattia F. 2016. A new empirical model for radar scattering from bare soil surfaces. Remote Sensing **8** : 1–14. DOI: 10.3390/rs8110920

Baghdadi N, Cresson R, El Hajj M, Ludwig R, La Jeunesse I. 2012. Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks. Hydrology and Earth System Sciences **16** : 1607–1621. DOI: 10.5194/hess-16-1607-2012

Baghdadi N, Hajj M El, Zribi M, Bousbih S. 2017. Calibration of the Water Cloud Model at C-Band for winter crop fields and grasslands. Remote Sensing **9** : 1–13. DOI: 10.3390/rs9090969

Balenzano A et al. 2021a. Sentinel-1 soil moisture at 1 km resolution: a validation study. Remote Sensing of Environment **263** DOI: 10.1016/j.rse.2021.112554

Balenzano A et al. 2021b. Sentinel-1 soil moisture at 1 km resolution: a validation study. Remote Sensing of Environment **263** DOI: 10.1016/j.rse.2021.112554

Balenzano A, Mattia F, Satalino G, Davidson MWJ. 2011a. Dense Temporal Series of C- and L-band SAR Data for Soil Moisture Retrieval Over Agricultural Crops. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **4** : 439–450. DOI: 10.1109/JSTARS.2010.2052916

Balenzano A, Mattia F, Satalino G, Davidson MWJ. 2011b. Use of dense temporal series of C-band SAR data for soil moisture retrieval over agricultural sites. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **4** : 439–450. DOI: 10.1109/JSTARS.2010.2052916

Balenzano A, Satalino G, Lovergine F, Rinaldi M, Iacobellis V, Mastronardi N, Mattia F. 2013. On the use of temporal series of L- and X-band SAR data for soil moisture retrieval. Capitanata plain case study. European Journal of Remote Sensing **46** : 721–737. DOI: 10.5721/EuJRS20134643

Bao Y, Lin L, Wu S, Kwal Deng KA, Petropoulos GP. 2018. Surface soil moisture retrievals over partially vegetated areas from the synergy of Sentinel-1 and Landsat 8 data using a modified water-cloud model. International Journal of Applied Earth Observation and Geoinformation **72** : 76–85. DOI: 10.1016/j.jag.2018.05.026 [online] Available from: https://doi.org/10.1016/j.jag.2018.05.026

Bargiel D. 2017. A new method for crop classification combining time series of radar

images and crop phenology information. Remote Sensing of Environment **198** : 369–383. DOI: 10.1016/j.rse.2017.06.022 [online] Available from: http://dx.doi.org/10.1016/j.rse.2017.06.022

Baronti S, Del Frate F, Paloscia S, Pampaloni P, Schiavon G, Solimini D. 1993. Sar polarimetric features of agricultural areas. 27–29 pp.

Bartalis Z, Scipal K, Wagner W. 2006. Azimuthal anisotropy of scatterometer measurements over land. IEEE Transactions on Geoscience and Remote Sensing **44** : 2083–2092. DOI: 10.1109/TGRS.2006.872084

Bauer-Marschallinger B et al. 2019. Toward Global Soil Moisture Monitoring with Sentinel-1: Harnessing Assets and Overcoming Obstacles. IEEE Transactions on Geoscience and Remote Sensing **57** : 520–539. DOI: 10.1109/TGRS.2018.2858004

Bauer-Marschallinger B et al. 2021. The normalised Sentinel-1 Global Backscatter Model, mapping Earth's land surface with C-band microwaves. Scientific Data **8** : 1–18. DOI: 10.1038/s41597-021-01059-7

Belgiu M, Csillik O. 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sensing of Environment **204** : 509–523. DOI: 10.1016/j.rse.2017.10.005 [online] Available from: https://doi.org/10.1016/j.rse.2017.10.005

Benninga HJF, van der Velde R, Su Z. 2022. Soil moisture content retrieval over meadows from Sentinel-1 and Sentinel-2 data using physically based scattering models. Remote Sensing of Environment **280** : 113191. DOI: 10.1016/j.rse.2022.113191 [online] Available from: https://doi.org/10.1016/j.rse.2022.113191

Berger M, Moreno J, Johannessen JA, Levelt PF, Hanssen RF. 2012. ESA's sentinel missions in support of Earth system science. Remote Sensing of Environment **120** : 84–90. DOI: 10.1016/j.rse.2011.07.023 [online] Available from: http://dx.doi.org/10.1016/j.rse.2011.07.023

Bindlish R, Barros AP. 2001. Parameterization of vegetation backscatter in radar-based, soil moisture estimation. Remote Sensing of Environment **76** : 130–137. DOI: 10.1016/S0034-4257(00)00200-5

BON. 1998. BON N.º 36 - 25/03/1998. BON N.º 36 - 25/03/1998 [online] Available from: http://www.lexnavarra.navarra.es/detalle.asp?r=28972.

Brocca L et al. 2011. Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. Remote Sensing of Environment **115** : 3390–3408. DOI: 10.1016/j.rse.2011.08.003 [online] Available from: http://dx.doi.org/10.1016/j.rse.2011.08.003

Brocca L, Tarpanelli A, Filippucci P, Dorigo W, Zaussinger F, Gruber A, Fernández-Prieto D. 2018. How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. International Journal of Applied Earth

Observation and Geoinformation **73** : 752–766. DOI: 10.1016/j.jag.2018.08.023 [online] Available from: https://doi.org/10.1016/j.jag.2018.08.023

Broersen PMT. 2006. Automatic Autocorrelation and Spectral Analysis . Springer

Brown RJ, Brisco BG, Edwards G. 1994. Adaptation of the MIMICS Backscattering Model to the Agricultural Context—Wheat and Canola at L and C Bands. IEEE Transactions on Geoscience and Remote Sensing **32** : 47–61. DOI: 10.1109/36.285188

Brown SCM, Quegan S, Morrison K, Bennett JC, Cookmartin G. 2003. High-resolution measurements of scattering in wheat canopies - Implications for crop parameter retrieval. IEEE Transactions on Geoscience and Remote Sensing **41** : 1602–1610. DOI: 10.1109/TGRS.2003.814132

Bruzzone L, Melgani F. 2005. Robust multiple estimator systems for the analysis of biophysical parameters from remotely sensed data. IEEE Transactions on Geoscience and Remote Sensing **43** : 159–173. DOI: 10.1109/TGRS.2004.839818

Burke M, Lobell D. 2010. Climate change and food security: adapting agriculture to a warmer world. Climate change and food security: Adapting agriculture to a warmer world : 201. DOI: 10.1007/978-90-481-2953-9 [online] Available from: http://rstb.royalsocietypublishing.org/content/360/1463/2139.short%5Cnhttp://link.springer.com/10.1007/978-90-481-2953-9

Canisius F, Shang J, Liu J, Huang X, Ma B, Jiao X, Geng X, Kovacs JM, Walters D. 2018. Tracking crop phenological development using multi-temporal polarimetric Radarsat-2 data. Remote Sensing of Environment **210** : 508–518. DOI: 10.1016/j.rse.2017.07.031 [online] Available from: https://doi.org/10.1016/j.rse.2017.07.031

Carranza C, Benninga H, van der Velde R, van der Ploeg M. 2019. Monitoring agricultural field trafficability using Sentinel-1. Agricultural Water Management **224** : 105698. DOI: 10.1016/j.agwat.2019.105698

Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. 2020. A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing **408** : 189–215. DOI: 10.1016/j.neucom.2019.10.118

Chai X, Zhang T, Shao Y, Gong H, Liu L, Xie K. 2015. Modeling and mapping soil moisture of plateau pasture using RADARSAT-2 imagery. Remote Sensing **7** : 1279–1299. DOI: 10.3390/rs70201279

Chan SK et al. 2016. Assessment of the SMAP Passive Soil Moisture Product. IEEE Transactions on Geoscience and Remote Sensing **54** : 4994–5007. DOI: 10.1109/TGRS.2016.2561938

Chen J, Jönsson P, Tamura M, Gu Z, Matsushita B, Eklundh L. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. Remote Sensing of Environment **91** : 332–344. DOI: 10.1016/j.rse.2004.03.014

Chen KL, Chen KS, Li ZL, Liu Y. 2015. Extension and validation of an advanced integral equation model for bistatic scattering from rough surfaces. Progress in Electromagnetics Research **152** : 59–76. DOI: 10.2528/PIER15011409

Chen L, Xing M, He B, Wang J, Shang J, Huang X, Xu M. 2021. Estimating Soil Moisture over Winter Wheat Fields during Growing Season Using Machine-Learning Methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **14** : 3706–3718. DOI: 10.1109/JSTARS.2021.3067890

Chuvieco E. 2019. Teledetección ambiental: La observación de la Tierra desde el Espacio . Ariel Ciencia

Chuvieco E, Huete A. 2009. Fundamentals of Satellite Remote Sensing . CRC Press: Boca Raton, FL, USA

d'Andrimont R, Verhegghen A, Lemoine G, Kempeneers P, Meroni M, van der Velde M. 2021. From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. Remote Sensing of Environment **266** DOI: 10.1016/j.rse.2021.112708

Dabrowska-Zielinska K, Musial J, Malinska A, Budzynska M, Gurdak R, Kiryla W, Bartold M, Grzybowski P. 2018. Soil moisture in the Biebrza Wetlands retrieved from Sentinel-1 imagery. Remote Sensing **10** DOI: 10.3390/rs10121979

Defourny P et al. 2019. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. Remote Sensing of Environment **221** : 551–568. DOI: 10.1016/j.rse.2018.11.007 [online] Available from: https://doi.org/10.1016/j.rse.2018.11.007

Dorigo W et al. 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sensing of Environment **203** : 185–215. DOI: 10.1016/j.rse.2017.07.001 [online] Available from: https://doi.org/10.1016/j.rse.2017.07.001

Dubois P, van Zyl J, Engman T. 1995. Measuring Soil Moisture with Imaging Radars. IEEE Transactions on Geoscience and Remote Sensing **33** : 915–926. DOI: 10.1109/TGRS.1995.477194

Entekhabi BD et al. 2015. TheSoilMoistureActive Passive (SMAP) Mission. Proceedings of the IEEE **98**

Entekhabi D, Reichle RH, Koster RD, Crow WT. 2010. Performance metrics for soil moisture retrievals and application requirements. Journal of Hydrometeorology **11** : 832–840. DOI: 10.1175/2010JHM1223.1

European Comission. 2021. List of potential agricultural practices that eco-schemes could support. : 5. [online] Available from: https://ec.europa.eu/info/sites/default/files/food-farming-

fisheries/key_policies/documents/factsheet-agri-practices-under-ecoscheme_en.pdf

FAO. 2018. The future of food and agriculture – Alternative pathways to 2050 [online] Available from: http://www.fao.org/3/I8429EN/i8429en.pdf

FAO. 2021. Food and Agriculture Organization of the United Nations. FAOSTAT. https://www.fao.org/faostat/en/#data/QCL

Feng Z, Zheng X, Li L, Li B, Chen S, Guo T, Wang X, Jiang T, Li X, Li X. 2021. Dynamic cosine method for normalizing incidence angle effect on C-band radar backscattering coefficient for maize canopies based on NDVI. Remote Sensing **13** DOI: 10.3390/rs13152856

Fieuzal R, Baup F, Marais-Sicre C. 2013. Monitoring wheat and rapeseed by using synchronous optical and radar satellite data—From temporal signatures to crop parameters estimation. Advances in Remote Sensing **02** : 162–180. DOI: 10.4236/ars.2013.22020

Foucras M, Zribi M, Albergel C, Baghdadi N, Calvet JC, Pellarin T. 2020. Estimating 500-m Resolution Soil Moisture Using Sentinel-1 and Optical Data Synergy. Water (Switzerland)

Fung AK. 1994. Microwave Scattering and Emission Models and their Applications . Artech House Publishers

Fung AK, Chen KS. 2004. An update on the IEM surface backscattering model. IEEE Geoscience and Remote Sensing Letters **1** : 75–77. DOI: 10.1109/LGRS.2004.826564

Gao B-C. 1996. NDWI - A normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space. Remote Sensing of Environment **58** : 257–266. DOI: 10.24059/olj.v23i3.1546

Gao Q, Zribi M, Escorihuela MJ, Baghdadi N, Segui PQ. 2018. Irrigation mapping using Sentinel-1 time series at field scale. Remote Sensing **10** : 1–18. DOI: 10.3390/rs10091495

Gauthier Y, Bernier M, Fortin JP. 1998. Aspect and incidence angle sensitivity in ers-1 sar data. International Journal of Remote Sensing **19** : 2001–2006. DOI: 10.1080/014311698215117

Gelman R, Carlin JB, Stern HS, Rubin DB. 1995. Bayesian Data Analysis . Input varibales: London

Gitelson A, Merzlyak MN. 1994. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. Journal of Plant Physiology **143** : 286–292. DOI: 10.1016/S0176-1617(11)81633-0

Graham AJ, Harris R. 2003. Extracting biophysical parameters from remotely sensed radar data: A review of the water cloud model. Progress in Physical Geography **27** :

217–229. DOI: 10.1191/0309133303pp378ra

Green JK, Seneviratne SI, Berg AM, Findell KL, Hagemann S, Lawrence DM, Gentine P. 2019. Large influence of soil moisture on long-term terrestrial carbon uptake. Nature **565** : 476–479. DOI: 10.1038/s41586-018-0848-x [online] Available from: http://dx.doi.org/10.1038/s41586-018-0848-x

Greifeneder F, Notarnicola C, Wagner W. 2021. A machine learning-based approach for surface soil moisture estimations with google earth engine. Remote Sensing **13** DOI: 10.3390/rs13112099

Hachani A, Ouessar M, Paloscia S, Santi E, Pettinato S. 2019. Soil moisture retrieval from Sentinel-1 acquisitions in an arid environment in Tunisia: application of Artificial Neural Networks techniques. International Journal of Remote Sensing **40** : 9159–9180. DOI: 10.1080/01431161.2019.1629503 [online] Available from: https://doi.org/10.1080/01431161.2019.1629503

Hajj M El, Baghdadi N, Zribi M, Bazzi H. 2017. Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas. Remote Sensing **9** : 1–28. DOI: 10.3390/rs9121292

El Hajj M, Baghdadi N, Zribi M, Belaud G, Cheviron B, Courault D, Charron F. 2016. Soil moisture retrieval over irrigated grassland using X-band SAR data. Remote Sensing of Environment **176** : 202–218. DOI: 10.1016/j.rse.2016.01.027 [online] Available from: http://dx.doi.org/10.1016/j.rse.2016.01.027

Hallikainen MT, Ulabz FT, Dobson MC, El-Rayes MA, Wu LK. 1985. Microwave Dielectric Behavior of Wet Soil-Part I: Empirical Models and Experimental Observations. IEEE Transactions on Geoscience and Remote Sensing **GE-23** : 25–34. DOI: 10.1109/TGRS.1985.289497

Han Y, Bai X, Shao W, Wang J. 2020. Retrieval of soil moisture by integrating Sentinel-1A and MODIS data over agricultural fields. Water (Switzerland) **12** DOI: 10.3390/w12061726

Harfenmeister K, Spengler D, Weltzien C. 2019. Analyzing temporal and spatial characteristics of crop parameters using Sentinel-1 backscatter data. Remote Sensing **11** : 1–30. DOI: 10.3390/rs11131569

He L, Qin Q, Panciera R, Tanase M, Walker JP, Hong Y. 2017. An extension of the alpha approximation method for soil moisture estimation using time-series sar data over bare soil surfaces. IEEE Geoscience and Remote Sensing Letters **14** : 1328–1332. DOI: 10.1109/LGRS.2017.2711006

Hoekman DH, Vissers AM. 2003. A new polarimetric classification approach evaluated for agricultural crops. IEEE Transactions on Geoscience and Remote Sensing **41** : 71–79. DOI: 10.1109/TGRS.2003.817795

Holtgrave AK, Förster M, Greifeneder F, Notarnicola C, Kleinschmit B. 2018. Estimation

of Soil Moisture in Vegetation-Covered Floodplains with Sentinel-1 SAR Data Using Support Vector Regression. PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science **86** : 85–101. DOI: 10.1007/s41064-018-0045-4 [online] Available from: https://doi.org/10.1007/s41064-018-0045-4

Hornáček M, Wagner W, Sabel D, Truong HL, Snoeij P, Hahmann T, Diedrich E, Doubková M. 2012. Potential for high resolution systematic global surface soil moisture retrieval via change detection using sentinel-1. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5** : 1303–1311. DOI: 10.1109/JSTARS.2012.2190136

Hütt C, Koppe W, Miao Y, Bareth G. 2016. Best accuracy land use/land cover (LULC) classification to derive crop types using multitemporal, multisensor, and multi-polarization SAR satellite images. Remote Sensing **8** DOI: 10.3390/rs8080684

Immitzer M, Vuolo F, Atzberger C. 2016. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote Sensing **8** DOI: 10.3390/rs8030166

Jackson TJ, Chen D, Cosh M, Li F, Anderson M, Walthall C, Doriaswamy P, Hunt ER. 2004. Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. Remote Sensing of Environment **92** : 475–482. DOI: 10.1016/j.rse.2003.10.021

Jagdhuber T, Hajnsek I, Bronstert A, Papathanassiou KP. 2013. Soil moisture estimation under low vegetation cover using a multi-angular polarimetric decomposition. IEEE Transactions on Geoscience and Remote Sensing **51** : 2201–2215. DOI: 10.1109/TGRS.2012.2209433

Jewell N. 1989. An evaluation of multi-date SPOT data for agriculture and land use mapping in the United Kingdom. International Journal of Remote Sensing **10** : 939–951. DOI: 10.1080/01431168908903936 [online] Available from: https://doi.org/10.1080/01431168908903936

Kaplan G, Fine L, Lukyanov V, Manivasagam VS, Tanny J, Rozenstein O. 2021. Normalizing the local incidence angle in sentinel-1 imagery to improve leaf area index, vegetation height, and crop coefficient estimations. Land **10** DOI: 10.3390/land10070680

Kerr YH et al. 2012. The SMOS soil moisture retrieval algorithm. IEEE Transactions on Geoscience and Remote Sensing **50** : 1384–1403. DOI: 10.1109/TGRS.2012.2184548

Kerr YH, Waldteufel P, Wigneron JP, Martinuzzi JM, Font J, Berger M. 2001. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Transactions on Geoscience and Remote Sensing **39** : 1729–1735. DOI: 10.1109/36.942551

Khabbazan S, Vermunt P, Steele-Dunne S, Ratering Arntz L, Marinetti C, van der Valk D, Iannini L, Molijn R, Westerdijk K, van der Sande C. 2019. Crop Monitoring Using Sentinel-1 Data: A Case Study from The Netherlands. Remote Sensing **11** : 1887. DOI:

10.3390/rs11161887

Koetz B, Defourny P, Bontemps S, Bajec K, Cara C, L DV, Kucera L, Malcorps P, Milcinski G. 2019. SEN4CAP Sentinels for CAP monitoring approach. 2019 JRC IACS workshop

Kong J, Yang J, Zhen P, Li J, Yang L. 2018. A coupling model for soil moisture retrieval in sparse vegetation covered areas based on microwave and optical remote sensing data. IEEE Transactions on Geoscience and Remote Sensing **56** : 7162–7173. DOI: 10.1109/TGRS.2018.2849009

Kornelsen KC, Coulibaly P. 2013. Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. Journal of Hydrology **476** : 460–489. DOI: 10.1016/j.jhydrol.2012.10.044 [online] Available from: http://dx.doi.org/10.1016/j.jhydrol.2012.10.044

Koster RD, Guo Z, Yang R, Dirmeyer PA, Mitchell K, Puma MJ. 2009. On the nature of soil moisture in land surface models. Journal of Climate **22** : 4322–4335. DOI: 10.1175/2009JCLI2832.1

Kussul N, Lemoine G, Gallego FJ, Skakun S V., Lavreniuk M, Shelestov AY. 2016. Parcel-based crop classification in Ukraine using Landsat-8 data and Sentinel-1A data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **9** : 2500–2508. DOI: 10.1109/JSTARS.2016.2560141

Lanchashire PD, Bleiholder H, Van Den Boom T, Langelüddeke P, Stauss R, Weber E, Witzenberger A. 1991. A uniform decimal code for growth stages of crops and weeds. Annals of Applied Biology **119** : 561–601. DOI: https://doi.org/10.1111/j.1744-7348.1991.tb04895.x [online] Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7348.1991.tb04895.x

Larrañaga A, Álvarez-Mozos J. 2016. On the added value of quad-pol data in a multi-temporal crop classification framework based on RADARSAT-2 imagery. Remote Sensing **8** : 1–19. DOI: 10.3390/rs8040335

Larranaga A, Alvarez-Mozos J, Albizua L, Peters J. 2013. Backscattering behavior of rain-fed crops along the growing season. IEEE Geoscience and Remote Sensing Letters **10** : 386–390. DOI: 10.1109/LGRS.2012.2205660

Lawrence H et al. 2014. Comparison between SMOS Vegetation Optical Depth products and MODIS vegetation indices over crop zones of the USA. Remote Sensing of Environment **140** : 396–406. DOI: 10.1016/j.rse.2013.07.021 [online] Available from: http://dx.doi.org/10.1016/j.rse.2013.07.021

Lee J Sen, Grunes MR, Pottier E. 2001. Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR. IEEE Transactions on Geoscience and Remote Sensing **39** : 2343–2351. DOI: 10.1109/36.964970

Lesiv M et al. 2019. Estimating the global distribution of field size using crowdsourcing.

Global Change Biology **25** : 174–186. DOI: 10.1111/gcb.14492

Li H, Zhang C, Zhang S, Atkinson PM. 2019. Full year crop monitoring and separability assessment with fully-polarimetric L-band UAVSAR: A case study in the Sacramento Valley, California. International Journal of Applied Earth Observation and Geoinformation **74** : 45–56. DOI: 10.1016/j.jag.2018.08.024 [online] Available from: https://doi.org/10.1016/j.jag.2018.08.024

Li J, Wang S. 2018. Using SAR-derived vegetation descriptors in a water cloud model to improve soil moisture retrieval. Remote Sensing **10** : 11–14. DOI: 10.3390/rs10091370

LIU C an, CHEN Z xin, SHAO Y, CHEN J song, Hasi T, PAN H zhu. 2019. Research advances of SAR remote sensing for agriculture applications: A review. Journal of Integrative Agriculture **18** : 506–525. DOI: 10.1016/S2095-3119(18)62016-7 [online] Available from: http://dx.doi.org/10.1016/S2095-3119(18)62016-7

Liu C, Chen Z, Shao Y, Chen J, Hasi T, Pan H. 2019. Research advances of SAR remote sensing for agriculture applications: A review. Journal of Integrative Agriculture **18** : 506–525. DOI: 10.1016/S2095-3119(18)62016-7 [online] Available from: http://dx.doi.org/10.1016/S2095-3119(18)62016-7

Liu C, Shang J, Vachon PW, McNairn H. 2013a. Multiyear crop monitoring using polarimetric RADARSAT-2 data. IEEE Transactions on Geoscience and Remote Sensing **51** : 2227–2240. DOI: 10.1109/TGRS.2012.2208649

Liu C, Shang J, Vachon PW, McNairn H. 2013b. Multiyear crop monitoring using polarimetric RADARSAT-2 data. IEEE Transactions on Geoscience and Remote Sensing **51** : 2227–2240. DOI: 10.1109/TGRS.2012.2208649

Liu L, Gudmundsson L, Hauser M, Qin D, Li S, Seneviratne SI. 2020. Soil moisture dominates dryness stress on ecosystem production globally. Nature Communications **11** : 1–9. DOI: 10.1038/s41467-020-18631-1 [online] Available from: http://dx.doi.org/10.1038/s41467-020-18631-1

Liu Y, Qian J, Yue H. 2021. Combined Sentinel-1A with Sentinel-2A to Estimate Soil Moisture in Farmland. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **14** : 1292–1310. DOI: 10.1109/JSTARS.2020.3043628

Lobell DB, Asner GP. 2004. Cropland distributions from temporal unmixing of MODIS data. Remote Sensing of Environment **93** : 412–422. DOI: 10.1016/j.rse.2004.08.002

Ma C, Li X, McCabe MF. 2020. Retrieval of high-resolution soil moisture through combination of Sentinel-1 and Sentinel-2 data. Remote Sensing **12** : 1–28. DOI: 10.3390/rs12142303

Macdonald HC, Waite WP. 1971. Soil Moisture Detection with Imaging Radars. Water Resources Research **7** : 100–110. DOI: 10.1029/WR007i001p00100

Macelloni G, Paloscia S, Pampaloni P, Marliani F, Gai M. 2001. The relationship between

the backscattering coefficient and the biomass of narrow and broad leaf crops. IEEE Transactions on Geoscience and Remote Sensing **39** : 873–884. DOI: 10.1109/36.917914

Matteson DS, James NA. 2014. A nonparametric approach for multiple change point analysis of multivariate data. Journal of the American Statistical Association **109** : 334–345. DOI: 10.1080/01621459.2013.849605

Mattia F, Le Toan T, Picard G, Posa FI, D'Alessio A, Notarnicola C, Gatti AM, Rinaldi M, Satalino G, Pasquariello G. 2003. Multitemporal C-band radar measurements on wheat fields. IEEE Transactions on Geoscience and Remote Sensing **41** : 1551–1560. DOI: 10.1109/TGRS.2003.813531

McNairn H, Brisco B. 2004. The application of C-band polarimetric SAR for agriculture: A review. Canadian Journal of Remote Sensing **30** : 525–542. DOI: 10.5589/m03-069

McNairn H, Champagne C, Shang J, Holmstrom D, Reichert G. 2009a. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. ISPRS Journal of Photogrammetry and Remote Sensing **64** : 434–449. DOI: 10.1016/j.isprsjprs.2008.07.006 [online] Available from: http://dx.doi.org/10.1016/j.isprsjprs.2008.07.006

McNairn H, Ellis J, Van Der Sanden JJ, Hirose T, Brown RJ. 2002. Providing crop information using RADARSAT-1 and satellite optical imagery. International Journal of Remote Sensing **23** : 851–870. DOI: 10.1080/01431160110070753

McNairn H, Shang J, Jiao X, Champagne C. 2009b. The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification. IEEE Transactions on Geoscience and Remote Sensing **47** : 3981–3992. DOI: 10.1109/TGRS.2009.2026052

Mercier A, Betbeder J, Baudry J, Le Roux V, Spicher F, Lacoux J, Roger D, Hubert-Moy L. 2020. Evaluation of Sentinel-1 & 2 time series for predicting wheat and rapeseed phenological stages. ISPRS Journal of Photogrammetry and Remote Sensing **163** : 231–256. DOI: 10.1016/j.isprsjprs.2020.03.009 [online] Available from: https://doi.org/10.1016/j.isprsjprs.2020.03.009

Meroni M, d'Andrimont R, Vrieling A, Fasbender D, Lemoine G, Rembold F, Seguini L, Verhegghen A. 2021. Comparing land surface phenology of major European crops as derived from SAR and multispectral data of Sentinel-1 and -2. Remote Sensing of Environment **253** DOI: 10.1016/j.rse.2020.112232

Militino AF, Moradi M, Ugarte MD. 2020. On the performances of trend and change-point detection methods for remote sensing data. Remote Sensing **12** : 1–25. DOI: 10.3390/rs12061008

Misra PN, Wheeler SG. 1978. Crop classification with LANDSAT multiespectral scanner data. Pattern Recognition **10** : 1–13.

Mladenova IE, Jackson TJ, Bindlish R, Hensley S. 2013. Incidence angle normalization of

radar backscatter data. IEEE Transactions on Geoscience and Remote Sensing **51** : 1791–1804. DOI: 10.1109/TGRS.2012.2205264

Modanesi S, Massari C, Bechtold M, Lievens H, Tarpanelli A, Brocca L, Zappa L, De Lannoy GJM. 2022. Challenges and benefits of quantifying irrigation through the assimilation of Sentinel-1 backscatter observations into Noah-MP. Hydrology and Earth System Sciences **26** : 4685–4706. DOI: 10.5194/hess-26-4685-2022

Moran S, Alonso L, Moreno JF, Cendrero Mateo MP, Fernando De La Cruz D, Montoro A. 2012. A RADARSAT-2 quad-polarized time series for monitoring crop and soil conditions in Barrax, Spain. IEEE Transactions on Geoscience and Remote Sensing **50** : 1057–1070. DOI: 10.1109/TGRS.2011.2166080

Mulla DJ. 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosystems Engineering **114** : 358–371. DOI: 10.1016/j.biosystemseng.2012.08.009 [online] Available from: http://dx.doi.org/10.1016/j.biosystemseng.2012.08.009

Naeimi V, Scipal K, Bartalis Z, Hasenauer S, Wagner W. 2009. An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. IEEE Transactions on Geoscience and Remote Sensing **47** : 1999–2013. DOI: 10.1109/TGRS.2008.2011617

Nasrallah A, Baghdadi N, El Hajj M, Darwish T, Belhouchette H, Faour G, Darwich S, Mhawej M. 2019. Sentinel-1 data for winter wheat phenology monitoring and mapping. Remote Sensing **11** DOI: 10.3390/rs11192228

Nguyen DB, Wagner W. 2017. European rice cropland mapping with Sentinel-1 data: The mediterranean region case study. Water (Switzerland) **9** : 1–21. DOI: 10.3390/w9060392

Notarnicola C. 2014. A Bayesian change detection approach for retrieval of soil moisture variations under different roughness conditions. IEEE Geoscience and Remote Sensing Letters **11** : 414–418. DOI: 10.1109/LGRS.2013.2264159

Notarnicola C, Angiulli M, Posa F. 2006. Use of Radar and Optical Remotely Sensed Data for Soil Moisture Retrieval Over Vegetated Areas. Ieee Transactions on Geoscience and Remote Sensing **44** : 925–935.

Notarnicola C, Angiulli M, Posa F. 2008. Soil Moisture Retrieval From Remotely Sensed Data: Neural Network Approach Versus Bayesian Method. IEEE Transactions on Geoscience and Remote Sensing **46** : 547–557. DOI: 10.1109/TGRS.2007.909951

Notarnicola C, Posa F. 2007. Inferring vegetation water content from C- and L-Band SAR Images. IEEE Transactions on Geoscience and Remote Sensing **45** : 3165–3171. DOI: 10.1109/TGRS.2007.903698

Ochsner TE et al. 2013. State of the art in large-scale soil moisture monitoring. Soil Science Society of America Journal **77** : 1888–1919. DOI: 10.2136/sssaj2013.03.0093

Oh Y. 2004. Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces. IEEE Transactions on Geoscience and Remote Sensing **42** : 596–601. DOI: 10.1109/TGRS.2003.821065

Oh Y, Sarabandi K, Ulaby FT. 1992. An Empirical Model and an Inversion Technique for Radar Scattering from Bare Soil Surfaces. IEEE Transactions on Geoscience and Remote Sensing **30** : 370–381. DOI: 10.1109/36.134086

Olofsson P, Foody GM, Herold M, Stehman S V., Woodcock CE, Wulder MA. 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment **148** : 42–57. DOI: 10.1016/j.rse.2014.02.015 [online] Available from: http://dx.doi.org/10.1016/j.rse.2014.02.015

Orynbaikyzy A, Gessner U, Conrad C. 2019. Crop type classification using a combination of optical and radar remote sensing data: a review. International Journal of Remote Sensing **40** : 6553–6595. DOI: 10.1080/01431161.2019.1569791 [online] Available from: https://doi.org/10.1080/01431161.2019.1569791

Ouaadi N et al. 2021. C-band radar data and in situ measurements for the monitoring of wheat crops in a semi-Arid area (center of Morocco). Earth System Science Data **13** : 3707–3731. DOI: 10.5194/essd-13-3707-2021

Ouaadi N, Jarlan L, Ezzahar J, Zribi M, Khabba S, Bouras E, Bousbih S, Frison PL. 2020. Monitoring of wheat crops using the backscattering coefficient and the interferometric coherence derived from Sentinel-1 in semi-arid areas. Remote Sensing of Environment **251** : 112050. DOI: 10.1016/j.rse.2020.112050 [online] Available from: https://doi.org/10.1016/j.rse.2020.112050

Ouellette JD et al. 2017. A Time-Series Approach to Estimating Soil Moisture from Vegetated Surfaces Using L-Band Radar Backscatter. IEEE Transactions on Geoscience and Remote Sensing **55** : 3186–3193. DOI: 10.1109/TGRS.2017.2663768

Le Page M, Jarlan L, El Hajj MM, Zribi M, Baghdadi N, Boone A. 2020. Potential for the detection of irrigation events on maize plots using Sentinel-1 soil moisture products. Remote Sensing **12** : 1–22. DOI: 10.3390/rs12101621

Palmisano D, Balenzano A, Satalino G, Mattia F, Pierdicca N, Monti-guarnieri A. 2018. Sentinel-1 sensitivity to soil moisture at high incidence angle and its impact on retrieval. 1430–1433 pp.

Pasolli L, Member S, Notarnicola C, Bruzzone L. 2011. Estimating Soil Moisture With the Support Vector Regression Technique. IEEE Geoscience and Remote Sensing Letters **8** : 1080–1084.

Pasolli L, Notarnicola C, Bertoldi G, Bruzzone L, Remelgado R, Greifeneder F, Niedrist G, Della Chiesa S, Tappeiner U, Zebisch M. 2015a. Estimation of soil moisture in mountain areas using SVR technique applied to multiscale active radar images at C-band. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **8** : 262–283. DOI: 10.1109/JSTARS.2014.2378795

Pasolli L, Notarnicola C, Bertoldi G, Bruzzone L, Remelgado R, Greifeneder F, Niedrist G, Della Chiesa S, Tappeiner U, Zebisch M. 2015b. Estimation of soil moisture in mountain areas using SVR technique applied to multiscale active radar images at C-band. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **8** : 262–283. DOI: 10.1109/JSTARS.2014.2378795

Peng J et al. 2021. A roadmap for high-resolution satellite soil moisture applications – confronting product characteristics with user requirements. Remote Sensing of Environment **252** : 112162. DOI: 10.1016/j.rse.2020.112162 [online] Available from: https://doi.org/10.1016/j.rse.2020.112162

Petitjean F, Inglada J, Gançarski P. 2012. Satellite image time series analysis under time warping. IEEE Transactions on Geoscience and Remote Sensing **50** : 3081–3095. DOI: 10.1109/TGRS.2011.2179050

Phung H-P, Nguyen L-D, Thong N-H, Thuy L-T, Apan AA. 2020. Monitoring rice growth status in the Mekong Delta, Vietnam using multitemporal Sentinel-1 data. Journal of Applied Remote Sensing **14** : 1. DOI: 10.1117/1.jrs.14.014518

Powell MJ. 1970. A new algorithm for unconstrained optimization. Nonlinear Programming : 31–66.

Pulvirenti L, Squicciarino G, Cenci L, Boni G, Pierdicca N, Chini M, Versace C, Campanella P. 2018. A surface soil moisture mapping service at national (Italian) scale based on Sentinel-1 data. Environmental Modelling and Software **102** : 13–28. DOI: 10.1016/j.envsoft.2017.12.022 [online] Available from: https://doi.org/10.1016/j.envsoft.2017.12.022

Qiu J, Crow WT, Wagner W, Zhao T. 2019. Effect of vegetation index choice on soil moisture retrievals via the synergistic use of synthetic aperture radar and optical remote sensing. International Journal of Applied Earth Observation and Geoinformation **80** : 47–57. DOI: 10.1016/j.jag.2019.03.015 [online] Available from: https://doi.org/10.1016/j.jag.2019.03.015

Rawls WJ, Brakensiek CL, Saxton KE. 1982. Estimation of soil water properties. Transactions of the American Society of Agricultural Engineers **25** DOI: 10.13031/2013.33720

Rizzoli P, Bräutigam B. 2014. Radar backscatter modeling based on global TanDEM-X mission data. IEEE Transactions on Geoscience and Remote Sensing **52** : 5974–5988.

Rouse JW, Haas RH, Schell JA, Deering DW. 1974. Monitoring vegetation systems in the Great Plains with ERTS. 309–317 pp.

Savitzky A, Golay MJE. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry **36** : 1627–1639. DOI: 10.1021/ac60214a048

Saxton KE, Rawls WJ. 2006. Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. Soil Science Society of America Journal **70** : 1569–

1578. DOI: 10.2136/sssaj2005.0117

Schaufler S, Bauer-Marschallinger B, Hochstöger S, Wagner W. 2018. Modelling and correcting azimuthal anisotropy in sentinel-1 backscatter data. Remote Sensing Letters **9** : 799–808. DOI: 10.1080/2150704X.2018.1480071 [online] Available from: https://doi.org/10.1080/2150704X.2018.1480071

Schlund M, Erasmi S. 2020. Sentinel-1 time series data for monitoring the phenology of winter wheat. Remote Sensing of Environment **246** : 111814. DOI: 10.1016/j.rse.2020.111814 [online] Available from: https://doi.org/10.1016/j.rse.2020.111814

Schmedtmann J, Campagnolo ML. 2015. Reliable crop identification with satellite imagery in the context of Common Agriculture Policy subsidy control. Remote Sensing **7** : 9325–9346. DOI: 10.3390/rs70709325

Schmedtmann J, Campagnolo ML, Thenkabail PS. 2015. Reliable crop identification with satellite imagery in the context of Common Agriculture Policy subsidy control. Remote Sensing **7** : 9325–9346. DOI: 10.3390/rs70709325

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ. 2010. Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Science Reviews **99** : 125–161. DOI: 10.1016/j.earscirev.2010.02.004 [online] Available from: http://dx.doi.org/10.1016/j.earscirev.2010.02.004

Sharma S, Swayne DA, Obimbo C. 2016. Trend analysis and change point techniques: a survey. Energy, Ecology and Environment **1** : 123–130. DOI: 10.1007/s40974-016-0011-1

Shi H, Lopez-Sanchez JM, Yang J, Li P, Zhao L, Zhao J. 2021. Contribution of Polarimetry and Multi-Incidence to Soil Moisture Estimation over Agricultural Fields Based on Time Series of L-Band SAR Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **14** : 300–313. DOI: 10.1109/JSTARS.2020.3036732

Shi J, Wang J, Hsu AY, O'Neill PE, Engman ET. 1997. Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data. IEEE Transactions on Geoscience and Remote Sensing **35** : 1254–1266. DOI: 10.1109/36.628792

Sitokonstantinou V, Papoutsis I, Kontoes C, Arnal AL, Andrés APA, Zurbano JAG. 2018. Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy. Remote Sensing **10** DOI: 10.3390/rs10060911

Skriver H. 2012. Crop classification by multitemporal C- and L-band single- and dual-polarization and fully polarimetric SAR. IEEE Transactions on Geoscience and Remote Sensing **50** : 2138–2149. DOI: 10.1109/TGRS.2011.2172994

Skriver H, Mattia F, Satalino G, Balenzano A, Pauwels VRN, Verhoest NEC, Davidson M.

2011. Crop classification using short-revisit multitemporal SAR data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **4** : 423–431. DOI: 10.1109/JSTARS.2011.2106198

Skriver H, Thoug M, G A. 1999a. Multitemporal C- and L-band polarimetric signatures of crops. IEEE Transactions on Geoscience and Remote Sensing **37** : 2413–2429.

Skriver H, Thoug M, G A. 1999b. Multitemporal C- and L-Band Polarimetric Signatures of Crops. IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING **37** : 2413–2429.

Small D. 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. IEEE Transactions on Geoscience and Remote Sensing **49** : 3081–3093. DOI: 10.1109/TGRS.2011.2120616

Song Y, Wang J. 2019. Mapping winter wheat planting area and monitoring its phenology using Sentinel-1 backscatter time series. Remote Sensing **11** DOI: 10.3390/rs11040449

Steele-Dunne SC, McNairn H, Monsivais-Huertero A, Judge J, Liu PW, Papathanassiou K. 2017. Radar remote rensing of agricultural canopies: A review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **10** : 2249–2273. DOI: 10.1109/JSTARS.2016.2639043

Stiles JM, Sarabandi K, Ulaby FT. 2000. Electromagnetic scattering from grassland-part II: measurement and modeling results. IEEE Transactions on Geoscience and Remote Sensing **38** : 349–356. DOI: 10.1109/36.823930

Torres R et al. 2012. GMES Sentinel-1 mission. Remote Sensing of Environment **120** : 9–24. DOI: 10.1016/j.rse.2011.05.028 [online] Available from: http://dx.doi.org/10.1016/j.rse.2011.05.028

Van Tricht K, Gobin A, Gilliams S, Piccard I. 2018. Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: a case study for Belgium. Remote Sensing : 1–22. DOI: 10.20944/preprints201808.0066.v1

Ulaby FT, Moore RK, Fung AK. 1982. Microwave Remote Sensing: Active and Passive . Vol II. Reading, MA: Addison-Wesley: Boston, MA

Ulaby FT, Moore RK, Fung AK. n.d. 'Microwave remote sensing: Volumes I, II and III . Artech House Inc., Dedham, Massachusetts

Ulaby FT, Sarabandi K, McDonald K, Whitt M, Dobson MC. 1990. Michigan microwave canopy scattering model. International Journal of Remote Sensing **11** : 1223–1253.

Vapnik VN. 1995. The Nature of Statistical Learning Theory . Springer-Verlag: New York

Veloso A, Mermoz S, Bouvet A, Le Toan T, Planells M, Dejoux JF, Ceschia E. 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sensing of Environment **199** : 415–426. DOI:

10.1016/j.rse.2017.07.015

Vereecken H, Huisman JA, Bogena H, Vanderborght J, Vrugt JA, Hopmans JW. 2010. On the value of soil moisture measurements in vadose zone hydrology: A review. Water Resources Research **46** : 1–21. DOI: 10.1029/2008WR006829

Verhoest NEC, Lievens H, Wagner W, Álvarez-Mozos J, Moran MS, Mattia F. 2008. On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. Sensors **8** : 4213–4248. DOI: 10.3390/s8074213

Verstraeten WW, Veroustraete F, Van Der Sande CJ, Grootaers I, Feyen J. 2006. Soil moisture retrieval using thermal inertia, determined with visible and thermal spaceborne data, validated for European forests. Remote Sensing of Environment **101** : 299–314. DOI: 10.1016/j.rse.2005.12.016

Wagner W, Blöschl G, Pampaloni P, Calvet JC, Bizzarri B, Wigneron JP, Kerr Y. 2007. Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. Hidrology research **38** : 1–20. DOI: 10.2166/nh.2007.029

Wagner W, Lemoine G, Rott H. 1999. A method for estimating soil moisture from ERS Scatterometer and soil data. Remote Sensing of Environment **70** : 191–207. DOI: 10.1016/S0034-4257(99)00036-X

Wang L, Qu JJ. 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. Geophysical Research Letters **34** : 1–5. DOI: 10.1029/2007GL031021

Wang Q, Li J, Jin T, Chang X, Zhu Y, Li Y, Sun J, Li D. 2020. Comparative Analysis of Landsat-8, Sentinel-2, and GF-1 Data for Retrieving Soil Moisture over Wheat Farmlands. Remote Sensing **12** : 1–16. DOI: https://doi.org/10.3390/rs12172708

Wang Z, Zhao T, Qiu J, Zhao X, Li R, Wang S. 2021. Microwave-based vegetation descriptors in the parameterization of water cloud model at L-band for soil moisture retrieval over croplands. GIScience and Remote Sensing **58** : 48–67. DOI: 10.1080/15481603.2020.1857123 [online] Available from: https://doi.org/10.1080/15481603.2020.1857123

Wardlow BD, Egbert SL, Kastens JH. 2007. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. Remote Sensing of Environment **108** : 290–310. DOI: 10.1016/j.rse.2006.11.021

Wasko C, Nathan R. 2019. Influence of changes in rainfall and soil moisture on trends in flooding. Journal of Hydrology **575** : 432–441. DOI: 10.1016/j.jhydrol.2019.05.054 [online] Available from: https://doi.org/10.1016/j.jhydrol.2019.05.054

Weiß T, Ramsauer T, Jagdhuber T, Löw A, Marzahn P. 2021. Sentinel-1 backscatter analysis and radiative transfer modeling of dense winter wheat time series. Remote Sensing **13** : 1–25. DOI: 10.3390/rs13122320

Wheeler SG, Misra PN. 1980. Crop classification with landsat multispectral scanner data II. Pattern Recognition **12** : 219–228. DOI: 10.1016/0031-3203(80)90061-8

Whelen T, Siqueira P. 2017. Use of time-series L-band UAVSAR data for the classification of agricultural fields in the San Joaquin Valley. Remote Sensing of Environment **193** : 216–224. DOI: 10.1016/j.rse.2017.03.014

Whelen T, Siqueira P. 2018. Time-series classification of Sentinel-1 agricultural data over North Dakota. Remote Sensing Letters **9** : 411–420. DOI: 10.1080/2150704X.2018.1430393 [online] Available from: https://doi.org/10.1080/2150704X.2018.1430393

Whitcraft AK, Vermote EF, Becker-reshef I, Justice CO. 2015. Cloud cover throughout the agricultural growing season : Impacts on passive optical earth observations. Remote Sensing of Environment **156** : 438–447. DOI: 10.1016/j.rse.2014.10.009 [online] Available from: http://dx.doi.org/10.1016/j.rse.2014.10.009

White E V., Roy DP. 2015. A contemporary decennial examination of changing agricultural field sizes using Landsat time series data. Geo: Geography and Environment **2** : 33–54. DOI: 10.1002/geo2.4

Xing M, He B, Ni X, Wang J, An G, Shang J, Huang X. 2019. Retrieving surface soil moisture over wheat and soybean fields during growing season using modified water cloud model from RADARSAT-2 SAR data. Remote Sensing **11** DOI: 10.3390/rs11161956

Xu L, Zhang H, Wang C, Zhang B, Liu M. 2019. Crop classification based on temporal information using Sentinel-1 SAR time-series data. Remote Sensing **11** : 1–18. DOI: 10.3390/rs11010053

Ye N, Walker JP, Rüdiger C. 2015. A Cumulative Distribution Function Method for Normalizing Variable-Angle Microwave Observations. IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING **53** : 3906–3916.

Zappa L, Schlaffer S, Brocca L, Vreugdenhil M, Nendel C, Dorigo W. 2022. How accurately can we retrieve irrigation timing and water amounts from (satellite) soil moisture? International Journal of Applied Earth Observation and Geoinformation **113** : 102979. DOI: 10.1016/j.jag.2022.102979 [online] Available from: https://doi.org/10.1016/j.jag.2022.102979

Zhang M, Lang F, Zheng N. 2021. Soil moisture retrieval during the wheat growth cycle using sar and optical satellite data. Water (Switzerland) **13** : 1–19. DOI: 10.3390/w13020135

Zhang X, Tang X, Gao X, Zhao H. 2018. Multitemporal Soil Moisture Retrieval over Bare Agricultural Areas by Means of Alpha Model with Multisensor SAR Data. Advances in Meteorology **2018** DOI: 10.1155/2018/7914581

Zhu L, Si R, Shen X, Walker JP. 2022. An advanced change detection method for time-series soil moisture retrieval from Sentinel-1. Remote Sensing of Environment **279** :

113137. DOI: 10.1016/j.rse.2022.113137 [online] Available from: https://doi.org/10.1016/j.rse.2022.113137

Zhuo W, Huang J, Li L, Zhang X, Ma H, Gao X, Huang H, Xu B, Xiao X. 2019. Assimilating Soil Moisture Retrieved from Sentinel-1 and Sentinel-2 Data into WOFOST Model to Improve Winter Wheat Yield Estimation. Remote Sensing **11** : 1618. DOI: 10.3390/rs11131618

Zribi M, Kotti F, Amri R, Wagner W, Shabou M, Lili-Chabaane Z, Baghdadi N. 2014. Soil moisture mapping in a semiarid region, based on ASAR/Wide Swath satellite data. Water Resources Research DOI: 10.1111/j.1752-1688.1969.tb04897.x