Charles University

Faculty of Science

Department of Applied Geoinformatics and Cartography

BSc. Júlia Šušková

Dynamic correction of vegetation effects in soil moisture modeling from SAR data using a change detection model

Dynamická korekce vlivu vegetace při modelování vlhkosti půdy z dat SAR pomocí modelu detekce změn

MASTER THESIS

Supervisor of the master thesis: Ing. Lukáš Brodský, Ph.D.

Study programme: Geoinformatics, Cartography and Remote Sensing

Prague 2024

Katedra aplikované geoinformatiky a kartografie

Přírodovědecká fakulta Univerzity Karlovy

Zadání diplomové práce

Jméno a příjmení: Júlia Šušková

Studijní program: Geoinformatika, kartografie a DPZ

Název: Korekce vlivu vegetace na odhad půdní vlhkosti z dat Sentinel-1 SAR

Zásady pro vypracování

Tématem této práce je stanovení vlhkosti půdy z dat Sentinel-1 SAR s korekcí na vliv vegetace. Půdní vlhkost je jednou ze základních klimatických veličin a zdokonalování modelů pro její určování bylo a je předmětem vědeckých studií. V této práci budou použita radarová data Sentinel-1 SAR získaná pomocí platformy GEE. Tato data spolu s databází ISMN (*International Soil Moisture Network*) budou použita k návrhu prototypu pro stanovení vlhkosti půdy. Dále bude navržen postup kalibrace a kalibrace modelu půdní vlhkosti z dat ISMN pro nejméně 3 vegetační období. Hlavním úkolem bude navrhnout korekci vlivu vegetace pomocí časových řad SAR CR (*Cross-polarization ratio*), ověřit výsledky a vyhodnotit přínos navržené korekce při modelování půdní vlhkosti z dat Sentinel-1 SAR.

V práci budou využity časové řady Sentinel-1 SAR s rozlišením přibližně 10 dní a bodová databáze ISMN s denními měřeními povrchové vlhkosti půdy. Pro stanovení vlhkosti půdy se použije metoda detekce změn. Časová řada CR bude použita ke korekci vlivu vegetace úpravou suché referenční hodnoty. Výstupem práce bude návrh, ověření a vyhodnocení metody korekce vlivu vegetace pro stanovení vlhkosti půdy z dat Sentinel.

Seznam odborné literatury:

B. Bauer-Marschallinger et al., "Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 1, pp. 520-539, Jan. 2019, doi: 10.1109/TGRS.2018.2858004.

M. Vreugdenhil, W. A. Dorigo, W. Wagner, R. A. M. de Jeu, S. Hahn and M. J. E. van Marle, "Analyzing the Vegetation Parameterization in the TU-Wien ASCAT Soil Moisture Retrieval," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 6, pp. 3513-3531, June 2016, doi: 10.1109/TGRS.2016.2519842.

Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sens. Environ. 2017, 199, 415–426.

Vedoucí diplomové práce: Ing. Lukáš Brodský, Ph.D.

Konzultant diplomové práce: Mariette Vreugdenhil

Datum zadání diplomové práce: 04.03.2023

doc. RNDr. Přemysl Štych, Ph.D.

....

Vedoucí diplomové práce   Garant studijního programu

V Praze dne

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

I declare that I have prepared my thesis independently and that I have included all information sources and literature used. Neither this thesis nor any substantial part of it has been has been submitted for another or the same academic degree.

In Prague date 30.07.2024   . . . .

                                                    Author's signature

Abstract: In times of climate change, soil moisture monitoring is an important aspect for its understanding and possibly mitigation. SAR data with high spatial resolution is an important tool for this purpose. This thesis deals with their use in SM retrieval from Sentinel-1 satellite data. The applied change detection model is further calibrated in order to remove the influence of vegetation on the resulting SM estimates by using the SAR variable cross-polarisation ratio. The RMSD decreased by 7% and the correlation increased by 8% using the calibration. The results presented do not achieve the accuracy of the ASCAT SM product but indicate the potential for vegetation correction using the Cross-polarization Ratio variable in further research to obtain a higher spatial resolution SM product.

Keywords: surface soil moisture, soil moisture retrieval, Sentinel-1, SM Change detection, vegetation correction

Abstrakt: V době klimatických změn je sledování půdní vlhkosti důležitým aspektem pro její pochopení. a případné zmírnění jejího dopadu. Data SAR s vysokým prostorovým rozlišením jsou pro tento účel důležitým nástrojem. Tento práce se zabývá jejich využitím při získávání SM z dat družice Sentinel-1. Použitý model detekce změn je dále kalibrován za účelem odstranění vlivu vegetace na výsledné odhady SM pomocí modelu SAR. proměnného poměru příčné polarizace. Kalibrací se RMSD snížila o 7% a korelace se zvýšila o 8%. Předložené výsledky nedosahují přesnosti produktu ASCAT SM, ale naznačují potenciál vegetace. korekce pomocí proměnné Cross-polarization Ratio v dalším výzkumu k získání produktu SM s vyšším prostorovým rozlišením.

Klíčová slova: povrchová vlhkost půdy, získávání vlhkosti půdy, Sentinel-1, detekce změn SM, korekce vegetace

Contents

List of Figures

List of Tables

List of Abbreviations

AIEM	Advanced Integral Equation Model
API	Application Programming Interface
ASCAT	Advanced SCATterometer
CLMS	Copernicus Land Monitoring Service
CR	Cross-polarization Ratio
DpRVIc	Dual-pol Radar Vegetation Index
ECV	Essential Climate Variable
EM	Electromagnetic spectrum
ENVISAT	Environmental Satellite mission
EOS	End of the Season
ERS	European Remote Sensing Satellite mission
ESA	European Space Agency
ESA CCI	European Space Agency's Climate Change Initiative
EUMETSAT	EU Organisation for the Exploitation of Meteorological Satellites
EVI	Enhanced Vegetation Index
GCOS	Global Climate Observing System
GEE	Google Earth Engine
GOM	Geometric Optics Model
GPI	Grid Point Index
GRD	Ground Range Detected
H-SAF	SAF on Support to Operational Hydrology and Water Management
HDA	Harmonized Data Access
HWSD	Harmonized World Soil Database
IEM	Integral Equation Model
ISMN	International Soil Moisture Network
IW	Interferometric Wide swath
L-BFGS-B	Limited-memory Broyden-Fletcher-Goldfarb-Shanno
LC	Land Cover
LIA	Local Incidence Angle
LST	Land Surface Temperature
LTDC	Long-Term Change Detection Model
MetOp	Meteorological Operational satellite mission

MIMICS	Michigan Microwave Canopy Scattering
MODIS	Moderate Resolution Imaging Spectroradiometer
MRS	Microwave Remote Sensing
NDVI	Normalized Difference Vegetation Index
NetCDF	Network Common Data Form
NIR	near-infrared EM (700–1400 nm)
OPTRAM	OPtical TRApezoid Model
POM	Physical Optics Model
PPI	Plant Phenology Index
RMSD	Root Mean Square Deviation
RS	Remote Sensing
RSMN	Romanian Soil Moisture Network
RTM	Radiative Transfer Model
S-1	Sentinel-1
S-2	Sentinel-2
SAF	Satellite Application Facility
SAR	Synthetic Aperture Radar
SM	Soil moisture
SNAP	Sentinel Application Platform
SPM	Small Perturbations Model
SOS	Start of the Season
SR	Surface Reflectance
STDC	Short-Term Change Detection Model
TB	brightness Temperature
TI	Thermal Inertia
TIR	thermal-infrared EM (NIR excluded) (1400 nm–1000 μm)
TxSON	Texas Soil Observation Network
UTC	Coordinated Universal Time
VIS	visible EM (380–700 nm)
VOD	Vegetation Optical Depth
VODCA	VOD Climate Archive
VPP	Vegetation Phenology Parameters
WARP	Soil Water Retrieval Package
WCM	Water Cloud Model

1 Introduction

Soil moisture is one of the 55 Essential Climate Variables (ECVs) defined by the Global Climate Observing System (GCOS). It stands for the water content stored in the vadose (unsaturated) soil zone. Soil moisture is an important variable for both the water and energy cycle [Seneviratne et al., 2010]. It plays a crucial role in vegetational evapotranspiration, which is of great importance as according to [Oki and Kanae, 2006], 60% of the total land precipitation returns to the atmosphere through this process. Additionally, land evapotranspiration consumes more than 50% of the entire solar energy absorbed by land surfaces [Trenberth et al., 2009]. These data confirm the great importance of soil moisture in monitoring climate change and the processes that influence it. Therefore, soil moisture is one of the variables that need to be monitored on a global scale and with sufficient temporal as well as spatial resolution. Ground measurements are still the most accurate method of determining soil moisture, but covering a larger area is particularly challenging. Hence, remote sensing (RS) methods are increasingly being explored for soil moisture monitoring purposes.

The presented thesis addresses the determination of surface soil moisture at field scale using the SM Change detection model and Sentinel-1 Synthetic Aperture Radar (SAR) data. Estimates from satellite sensors are compared with in-situ measurements from the International Soil Moisture Network database (ISMN) in Romania in order to evaluate the model performance. Furthermore, a model calibration is proposed in order to correct for vegetation influence on the SM estimate from the Sentinel-1 satellite. The results of the calibrated model are compared with the original Change detection model without vegetation correction, with in-situ ISMN data as well as with the established SM product from the Advanced SCATterometer sensor (ASCAT).

In the following Introduction section (1.1), different retrieval methods for soil moisture estimation using RS techniques are introduced in the scope of relevant studies within the past decades. Strong emphasis is put on Microwave RS (MRS) as it is still the preferred method to estimate surface SM from RS data and is the primary method utilized for currently available SM products at global or regional scale [Li et al., 2021]. In addition, different approaches to correct for vegetation impact are also introduced in the section 1.2. The 1.3 section introduces relevant principles of measuring the Earth's surface using SAR sensors, particularly the C-band SAR sensor on board of Sentinel-1. The last section (1.4) is dedicated to the Cross-polarization Ratio variable (CR) and its

importance in describing vegetation since CR is used in the current thesis for vegetation correction in the SM Change detection model.

1.1 Soil moisture retrieval methods from RS

Over the past 50 years, extensive research has been dedicated to generate reliable soil moisture products from RS techniques rather than from costly and time-expensive ground-based measurements [Liu and Yang, 2022]. A wide variety of space-borne sensors receiving reflected or emitted energy from the Earth's surface has been applied to monitor SM. Signal from visible (VIS), near-infrared (NIR), thermal-infrared (TIR) as well as microwave electromagnetic spectrum has been exploited to retrieve information about surface SM in the past [Petropoulos et al., 2015, Zhang and Zhou, 2016, Li et al., 2021].

1.1.1 Optical RS methods

Retrieving information about SM from VIS/NIR electromagnetic spectrum (350–1 400 nm) is based on the relationship between changes in surface soil reflectance properties [Babaeian et al., 2019]. Alterations of soil reflectance due to different water content have been observed already early in the 20th century [Ångström, 1925, Bowers and Smith, 1972]. A decrease in soil surface reflectance with an increase in soil water content can be observed due to energy absorption within the NIR band [Ångström, 1925]. Several empirical models exploiting either spectral information from a single band or combining spectral signal into an index have been proposed [Li et al., 2021]. The problem associated with all empirical models is the strong site dependency on specific surface conditions [Petropoulos et al., 2015, Li et al., 2021]. Albeit less frequently used, physically-based models attempt to overcome this obstacle [Li et al., 2021]. The OPtical TRApezoid Model (OPTRAM) proposed by [Sadeghi et al., 2017] exploits the linear relationship between surface SM and short-wave infrared reflectance. The results have shown a Root Mean Square Deviation ($RMSD$) of less than 0.05 $m^3.m^{-3}$ applied on Sentinel-1 and Landsat-8 data with high spatial resolution for VIS/NIR (10–30 m).

TIR electromagnetic signal (1 400–14 000 1.55 µm) measuring the land surface temperature (LST) differences can be utilized to obtain estimates of thermal inertia (TI) [Petropoulos et al., 2015]. TI is a function of thermal conductivity and heat capacity [Zhang and Zhou, 2016]. Water abundance due to increased SM causes a rise in both variables and in this way indirectly leads to an increase of TI [Zhang and Zhou, 2016].

Despite a very clear physical connection between SM and TI, which can be estimated from diurnal LST observations [Kahle, 1977], this method is not well suited for moderately or densely vegetated soils with an evapotranspiration effect [Li et al., 2021] and is also influenced by the soil type, therefore, it is not transferable and applicable at larger scales [Petropoulos et al., 2015].

1.1.2 Microwave active and passive RS methods

Microwave RS methods are preferred when retrieving SM [Barrett et al., 2009, Li et al., 2021]. It encompasses passive and active sensors measuring signal within the wavelength range of 1 mm–30 cm. For SM retrieval, signal from the X-, C- and L-band is generally used as it enables reliable data about soil dielectric properties for the first 0–1 cm (X-, C-band) and 0-5 cm (L-band) of the surface soil [Lv et al., 2018, Li et al., 2021]. They are widely applied also because of their low frequency which allows to neglect scattering or absorbing effects of atmosphere molecules [Karthikeyan et al., 2017]. Therefore, it is assumed that all interactions with any medium take place first at the surface-air boundary. L-band is broadly accepted as the best option for SM detection as C- and X-band signal might not penetrate through vegetation deeply enough [Petropoulos et al., 2015]. In addition, [Njoku and Entekhabi, 1996] showed that the L-band is less sensitive toward vegetation water content up to 5 $kg.m^2$ in comparison to the X- and C-band (1.5 $kg.m^2$). The main principle of moisture detection is built upon the difference between soil dielectric constant during wet and dry conditions. Consequently, SM retrieval models for microwave sensors exploit the relationship between sensor-measured quantity (surface emission by passive MRS and backscatter coefficient by active MRS) and dielectric properties of soil, and the dependency on the dielectric constant and SM [Karthikeyan et al., 2017].

Passive sensors measure surface thermal emission as brightness temperature TB, which can be indirectly related to soil water content [Karthikeyan et al., 2017]. An increase in SM reduces the emissivity of the soil surface which further leads to a decrease in TB [Karthikeyan et al., 2017]. In the first studies, several empirical models have been developed connecting measured TB with SM [Li et al., 2021]. These models were usually dependent on land cover or specific surface conditions. Therefore, physical approaches became more popular. They consist of a combination of two models – the Radiative Transfer Model (RTM) and the dielectric mixing model

[Karthikeyan et al., 2017]. As for the first one, through solving radiative transfer equations TB is related to dielectric constant under various conditions (e.g., temperature, surface roughness, vegetation, atmosphere, and cosmic background) and sensor characteristics [Karthikeyan et al., 2017]. Nowadays, the so-called t-ω model [Mo et al., 1982] is the common model used in the first retrieval phase and is applied in all passive MRS SM retrieval models [Babaeian et al., 2019]. It describes the interaction of microwave wavelength with the soil surface and vegetation. Thus, the contribution of the soil surface can be separated from that of vegetation and further inverted to soil moisture estimates. The model requires two input parameters – vegetation optical depth (VOD) as an indicator of radiation attenuation by vegetation and single scattering albedo ω describing the portion of radiation scattered by vegetation relatively to the total extinction [Li et al., 2021].

In the next step, when deriving the relation between SM and dielectric constant multiple dielectric mixing models have been proposed – Wang and Schmugge model [Wang and Schmugge, 1980], Dobson model [Dobson et al., 1985], Hallikainen model [Hallikainen et al., 1985] or Mironov model [Mironov et al., 2009]. The models vary in the input parameters besides SM (particle size, porosity, microwave frequency, dry soil bulk density, etc.) [Karthikeyan et al., 2017]. In the early years of the mission, the Dobson model was implemented within the retrieval algorithm until 2013 when it was supplemented by the Mironov model [Mialon et al., 2015].

The concept of SM derivation from active microwave signal is similarly based on the relation between soil dielectric properties and soil water content. The difference is that active microwave instruments transmit their radiation at a specific wavelength and measure the reflected backscatter coefficient. The backscatter coefficient σ^0 is defined as the radar cross section σ normalized over an area A. Active MRS allows to enhance the spatial resolution of the final data compared to passive microwave data since the natural surface emissions within the microwave electromagnetic spectrum are weak and therefore a large footprint is required to capture a sufficient amount of the signal resulting in coarse spatial resolution [Karthikeyan et al., 2017]. The active microwave signal can be absorbed, scattered, reflected, or further transmitted into the deeper medium layers [Kornelsen and Coulibaly, 2013]. Therefore, the sensor can only measure the fraction of the original signal scattered back in the sensor's direction. This fraction depends on surface conditions and sensor properties [Kornelsen and Coulibaly, 2013].

Decoupling of the total received backscatter signal to retrieve only information about soil water content is a complex issue affected by several factors such as incidence angle, SM, the contribution of different backscattering mechanisms (specular, intermediate specular/diffuse, diffuse scattering) as well as sensor transmitter/receiver polarization properties [Verhoest et al., 2008].

Like passive MRS, the backscatter coefficient must be defined as a function of SM or dielectric constant to be later inverted to model SM content. These models can be classified according to [Barrett et al., 2009] into physically-based, semi-empirical, and empirical models. The simplest model linking SM and the backscatter coefficient is a linear model [Ulaby et al., 1978]. As highlighted by [Verhoest et al., 2008] this model is not transferable between study sites as the slope of the linear fit varies strongly depending on the location. Moreover, [Barrett et al., 2009] emphasized that this model is designed for specific datasets and under specific sensor parameters.

Semi-empirical models unlike empirical models are designed upon a physical/theoretical basis but still require experimental data to be calibrated [Verhoest et al., 2008]. [Oh et al., 1992] simulated the CR of the backscatter coefficient in different polarizations as a function of SM and roughness conditions given by normalized root mean square (rms) surface roughness. The advantage of this model is the single parameter needed to account for surface roughness. Nevertheless, many studies proved its unsuitability for airborne and spaceborne SAR sensors [Boisvert et al., 1997, van Oevelen and Hoekman, 1999, Baghdadi and Zribi, 2006]. Dubois model designed by [Dubois et al., 1995] simulates the co-polarized backscatter coefficient as a function of soil dielectric constant and rms of surface roughness. The model was found to be less sensitive to system noise due to accounting only for co-polarized signal [Barrett et al., 2009].

Physically based models can be implemented when surface roughness parameters are known [Barrett et al., 2009]. Small Perturbations Model (SPM), Geometric Optics Model (GOM), Physical Optics Model (POM), or Integral Equation Model (IEM) which incorporates SPM are used to model the interaction of electromagnetic waves with rough surface. All of these models have their shortcomings and lead to better results under different roughness conditions [Barrett et al., 2009]. Overall, IEM and its modified version Advanced IEM (AIEM) are most widely applied as it is applicable across a broader surface roughness range [Verhoest et al., 2008, Barrett et al., 2009]. It should be noted that all of these models neglect the impact of vegetation cover and are therefore suitable for bare soil, in areas with sparse vegetation cover [Verhoest et al., 2008,

Barrett et al., 2009]. Additionally, in arid regions, IEM also neglects the sub-surface scattering effect [Verhoest et al., 2008].

The aforementioned models can be used to simulate backscatter from single imagery correcting for soil roughness or vegetation. They are the forward models of backscatter behavior under various environmental and sensor conditions. SM is further simulated using the dielectric mixing model as a function of measured backscatter. The parameterization of surface and vegetation properties is, however, difficult in most cases [Kornelsen and Coulibaly, 2013]. Therefore, Change detection models are a very popular SM retrieval method. They assume that surface roughness and land cover conditions stay constant over time. As stated in [Zhu et al., 2022] there are long-term and short-term Change detection models (LTCD, STDC). STDC requires a collection of temporally dense imageries and uses the time series to retrieve SM [Balenzano et al., 2021]. LTCD was originally designed by [Wagner et al., 1999] and assumes a linear relationship between observed backscatter and long-term SM range. SM range (*dynamic range* or *sensitivity*) is defined as the difference between long-term reference backscatter during the driest and wettest conditions. This method was initially developed for the former European Remote Sensing Satellite (ERS) mission and later re-used in the past Environmental Satellite (ENVISAT) mission and the current Meteorological Operational satellite (MetOp) mission carrying the ASCAT scatterometer. The backscatter references are found by applying the cross-angle concept which assumes that under a specific incidence angle, the backscatter signal is affected solely by SM. The angle reference was set empirically to 25° and 40° for dry and wet conditions, respectively [Naeimi et al., 2009]. The Change detection method is very simple and hence was applied to provide global SM products from previously mentioned spaceborne missions. Both STCD and LTDC were also applied to Sentinel-1 SAR data to generate a 1-km macroscale SM product [Bauer-Marschallinger et al., 2019, Balenzano et al., 2021]. The main difference within the applied methods is the character of retrieved SM. In the former study (applying LTCD) relative SM scaled between dry and soil reference is computed whereas in the latter one absolute values of SM were estimated. In addition, according to [Zhu et al., 2022], the constant soil roughness assumption can be found valid for the STCD but is vague for LTCD.

1.2 Vegetation correction

To simulate vegetation impact on the measured backscatter coefficient, models for vegetation cover were designed in previous research – Michigan microwave canopy scattering (MIMICS) and Water Cloud Model (WCM). The former model defines a set of very specific geometrical parameters to characterize forest canopy and divides it into three parts – crown, trunk, and ground [Ulaby et al., 1990]. Compared to MIMICS, WCM simplifies the structure of vegetation to minimize the amount of describing parameters. WCM assumes the vegetation layer to be represented by water droplets distributed randomly within a "cloud" layer of air and vegetation. The theory behind this concept is the assumption that vegetation water content is the main factor affecting vegetation backscatter and that up to 99% of vegetation volume consists of air [Attema and Ulaby, 1978]. With this model definition, backscatter can be simulated requiring only four parameters. Two of these parameters, extinction of microwave signal due to vegetation water content VWC and extinction due to vegetation scattering, are linked to the parameters t and ω in the previously mentioned $t - \omega$ model. In essence, WCM combines the $t - \omega$ model to account for vegetation backscatter and the linear model for backscatter over bare soil [Wagner et al., 2019]. One shortcoming of the WCM model is the neglect of multiple scattering effects [Kornelsen and Coulibaly, 2013]. In the latest version of WCM [Kweon and Oh, 2015] the authors included new parameters to characterize the angular distribution of leaves which led to lower $RMSD$ for backscatter in vertical-vertical (VV) and horizontal-horizontal (HH) polarization over corn and bean fields compared to the original WCM. According to [Ma et al., 2020], the MIMICS model is more suitable for forest canopy as it was originally developed for this land cover and only later modified for vegetation of lower height [Li et al., 2021]. On the other hand, it requires insertion of many parameters that are difficult to obtain and, hence, it is not as widely applied as the WCM [Li et al., 2021]. WCM is also implemented in Terra-Sar spaceborne-mission providing SM product at a global scale [Liu and Yang, 2022].

In the case of the SM product from the ASCAT sensor, the above-mentioned cross-angle principle can be modified so that the model also directly captures changes in vegetation and thus corrects for their effect on the resulting soil moisture estimate. In their studies [Pfeil et al., 2018, Hahn et al., 2021b, Vreugdenhil et al., 2016], researchers from the Vienna University of Technology (TU Wien) showed that the selection of reference cross-angles and the subsequent impact on the determination of dry and wet reference in the Change detection model helped to improve the soil moisture

determination.

1.3 SAR specifics

Synthetic Aperture Radar (SAR) belongs to the active RS. SAR sensors transmit microwave pulses with a wavelength between 1 mm and 1 m and measure the backscattered intensity of the transmitted puls as well as the time lag between transmission and signal reception. Depending on the wavelength of the emitted pulse, SAR sensors can be distinguished according to the different bands of the microwave spectrum, which are indicated by letter abbreviations from Ka $(0.75 - 1.1$ $cm)$ to P $(0.3 - 1$ $m)$. The most commonly used bands are X $(2.4 - 3.75$ $cm)$, C $(3.75 - 7.5$ $cm)$ and L $(15 - 30$ $cm)$. The longer wavelength signal penetrates deeper to the surface through terrestrial objects while the shorter wavelength signal can also record details of objects on the surface. Sentinel-1 carries a C-band sensor on its board.

1.3.1 Acquisition geometry

SAR sensors are characterized by their side-looking measuring technique. The signal is transmitted in an oblique direction to the surface of the earth. Depending on the direction, two resolutions are distinguished in SAR systems – across-track (*range*) and along-track (*azimuth*) resolution. SAR *range* depends on the length of the transmitted pulse. If the distance between two objects on the surface is greater than half of the pulse's length (in range dimension, not on the ground), the objects can be separated from each other by the sensor. The beamwidth determines the azimuthal resolution as with the increasing distance from the sensor the radar illumination also increases and, consequently, the resolution gets coarser. Accordingly, two objects in the so-called *near range* can be separable but might not be in the *far range.* The along-tack resolution can be enhanced considerably by increasing the antenna length since the beamwidth is inversely proportional to this length. With a longer antenna, the beamwidth gets narrower and therefore the along-track resolution is improved. SAR data are typical for their high spatial resolution, which is achieved using the synthetic aperture technique. SAR sensors use this technique to simulate a longer antenna by measuring not only the intensity of the received pulse but also its phase and by means of the Doppler effect the respective distance from a specific sensor position relatively to the Earth surface can be determined. This technique replaces the need for a long physical antenna that is impossible to install on the carrier.

Sentinel-1 mission is further characterized by a regular orbital path and a finite number of relative orbits over a specific location on Earth surface. In other words, a particular location is always observed by the C-SAR sensor on Sentinel-1 from two (at the equator) up to six (at the poles) positions. The regular acquisition geometry of C-SAR Sentinel-1 influences also the temporal analysis of SAR time series. Depending on the current relative orbit, the incidence angle θ of the signal, defined as the angle between the nadir and the direction of the SAR signal, varies. The change in θ directly influences the intensity of the measured signal back by the sensor. When using imageries from different relative orbits of a sensor, the difference in the incidence angle has to be considered and corrected to achieve comparability between them. The difference in the intensity of the signal from various orbits can be up to 2.7 dB as found by [Gauthier et al., 1998]. According to [Dobson and Ulaby, 1986], a low incidence angle of 10°–20° is recommended for SM retrieval for the C-band of an active microwave sensor. In the thesis, specifically, the impact of Local incidence angle (LIA) is corrected which is defined as the angle between normal to the Earth surface and the direction of the SAR signal to account also for the terrain properties at specific locations.

1.3.2   Polarization

Microwave wavelengths can be characterized by the orientation of their electromagnetic field. The polarization is defined by the direction of the electric component of the field, which is oriented always perpendicularly to the magnetic component. SAR sensors can transmit electromagnetic waves of a known polarization. The transmitted signal interacts with the objects on the ground. Depending on the object's properties the interaction can depolarize the transmitted signal which leads to a different polarization state of the signal traveling back to the antenna, where it is subsequently recorded. Sentinel-1 C-SAR can transmit and receive signal from different polarizations and supports the so-called dual-polarization. This means that the sensor transmits signal at one polarization (vertical or horizontal) and is able to record the backscattered signal at both polarizations (vertical+horizontal or horizonal+vertical). The dual polarization is mostly known and noted as VV+VH or HH+HV considering Sentinel-1. Measuring more than only one polarization can reveal more information about the complexity of the object's properties. Objects that are dominant more in vertical direction (e.g., tree stems, branches, buildings) lead to a stronger response for a vertically polarized wave. Surface structures of an irregular form can, on the contrary, lead to depolarization and

enhance the opposite polarization in the backscattered signal than in the transmitted one. When considering the soil moisture retrieval and the impact of polarization, HH polarization proved to be more accurate to the surface soil moisture than to the objects on the surface than HV [Barrett et al., 2009]. However, [Holah et al., 2005] found stronger sensitivity to SM in VV polarization than in HV or HH.

1.4 Cross-polarization Ratio

The total measured backscatter signal can be described as a sum of three scattering components [Vreugdenhil et al., 2020] – volume scattering, surface scattering, and interaction scattering. This fractional composition of the total measured backscatter coefficient changes during vegetation growth with increased contribution of vegetation-inflicted volume scattering and vegetation water content [Veloso et al., 2017]. In this study, the authors propose that CR (VH/VV) is a good indicator for monitoring vegetation growth being less sensitive towards SM or double-bouncing effects than single polarized backscatter. However, CR temporal behavior differs over various land cover types, particularly cropland/grassland and broadleaf deciduous forest. [Vreugdenhil et al., 2020] studied the relationship between CR from backscatter data of Sentinel-1 characterized by its high spatial resolution and two VOD datasets from passive radar measurements – the first one obtained from active microwave scatterometer ASCAT and the second one from the long-term VOD Climate Archive (VODCA). VOD was retrieved utilizing the combined TU Wien Change detection model/Water-Cloud-Model and Land Parameter Retrieval Model for active and passive radar data, respectively. CR and VOD over grassland and cropland correspond with each other the best, with a Pearson correlation coefficient greater than 0.75 over the entire Europe. Less satisfying correlations were found over needleleaf tree land cover, with a moderately low Pearson coefficient of 0.38. Surprising behavior of CR time series is found for areas covered by broadleaf deciduous forests. As CR is expected to be more sensitive to changes in vegetation rather than SM [Veloso et al., 2017, Khabbazan et al., 2019, Vreugdenhil et al., 2020], an increase in volume scattering during the tree leaf growth in spring/summer was anticipated to lead to a rise in CR as well. However, CR decreases in summer while VOD increases over broadleaf deciduous forests resulting in a negative correlation of -0.33 over entire Europe. During winter, a rise in CR is observed. The authors argue this might be caused by soil-vegetation interactions from branches and stems. CR decrease during summer is further addressed in a very recent study by [Yu et al., 2023] and explained

to be a consequence of strong vegetation attenuation and a decrease in woody water content as a result of leaf transpiration.

2 Data

This chapter presents the datasets utilized in the thesis, organized into five key sections. Section 2.1 introduces the in-situ data used for ground-truth validation. Section 2.2 describes the Sentinel-1 satellite data employed for the soil moisture retrieval model. Section 2.3 is subdivided into four parts. Subsection 2.3.1 describes the vegetation index NDVI, which is used for the comparison with the primary calibration variable CR. Subsection 2.3.2 introduces the Vegetation Phenology Parameters dataset (VPP) used to define seasonal periods in the time series. Subsection 2.3.3 describes the ASCAT SM product used to benchmark the results against an established soil moisture product. Finally, the subsection 2.3.4 includes ISMN in-situ precipitation data to analyze the model's capability to capture rainfall events, and ISMN soil temperature data used for filtering purposes.

2.1 ISMN Soil Moisture

The International Soil Moisture Network (ISMN) is a collaborative initiative fostering collecting, sharing, and distributing in-situ soil moisture data. It serves as a platform where researchers, scientists, and organizations from around the world can contribute their soil moisture measurements to improve our understanding of soil moisture dynamics across different regions and climates. The data is provided by the cooperating stations and can be measured at different sampling rates or units, across various time zones or in multiple depth levels. Therefore, some level of standardization is necessary to achieve a harmonized database [Dorigo et al., 2021]. The common standard sampling rate for the ISMN database is hourly Coordinated Universal Time (UTC) reference time steps. Therefore, in the provided data the closest measurements within the 0.5-hour interval are always matched with the reference [Dorigo et al., 2021].

The data is available on the ISMN website, where it is possible to filter the data by location, time period, properties of the measuring sensors and the surrounding environmental conditions of the stations. Each ISMN dataset is distributed with mandatory metadata originating from an external source [Dorigo et al., 2021]. The mandatory metadata contains the respective Köppen-Geiger category (source data resolution is 0.1°) [Peel et al., 2007], land cover (LC) information for years 2000, 2005 and 2010 from the European Space Agency's Climate Change Initiative (ESA CCI) (version 1.6.1 with 300 m resolution) and soil properties from the Harmonized World Soil Database (HWSD, ver-
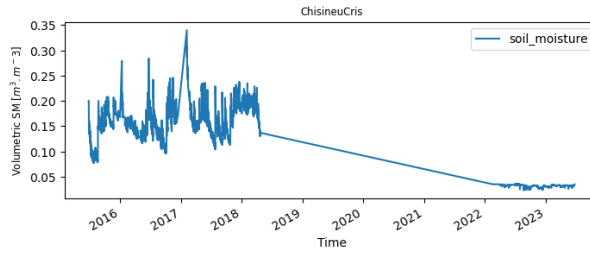
Figure 2.1 Low plateau, data source: *ISMN*

sion 1.1 with 1 km resolution) provided by [FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009].

Soil moisture data is accessible via the Python package *ismn* developed and published by [Preimesberger et al., 2023]. The ISMN database can be further converted into a *pandas* dataframe directly in a Python interpreter using the *pandas* library, which enables an efficient data handling via data structures like DataFrame and Series[1].

2.1.1 Site description and pre-processing

For the master thesis, data from the Romanian Soil Moisture Network (RSMN) incorporated in the ISMN initiative was used. RSMN encompasses together 20 measuring stations distributed across the entire country. The measurements are harmonized using the same technique and are available for the depth of 0–0.05 m (surface soil layer). Currently, not all of the stations provide soil moisture measurements (e.g., *Slobozia*, *Adamclisi*). In addition, as reported on the official website of ISMN one of the stations (*Chisineu Cris*) shows characteristics of "low plateau" since April 2018 (Fig. 2.1)[2]. The measuring sites are located primarily within the borders of meteorological stations or directly next to them. The majority of the RSMN stations are located on grassland. Cropland is often found in their immediate vicinity. However, multiple RSMN stations are surrounded by high objects such as agricultural/industrial buildings or residences and their presence might impact the radar signal. Therefore, due to short SM measurements archive, dubious SM behavior and presence of high objects capable of disrupting SAR signal, only 10 stations were eventually selected for further work – *Banloc, Corugea, Cotnari, Dej, Dumbraveni, Oradea, RosioriideVede, SannicolauMare, SatuMare, Tecuci.* The aerial imageries for the selected stations are displayed in Figure 2.2 and they are listed in Table 4.1.

Before the in-situ RSMN data was merged with the satellite data, measurements

---

[1] https://pandas.pydata.org/docs/
[2] https://ismn.earth/en/news/

not labeled as *good (G)* were eliminated from the dataset. The ISMN database provides the quality labels together with the SM measurements for each data point [Dorigo et al., 2021]. Table 4.1 lists for each station the assigned Köppen-Geiger, ESA CCI LC class, time window covering SM measurements and the fraction of reliable measurements of the original dataset. Reliable measurements preserved for further analysis represent at least 71% of the original dataset. It is worth noting, that even though the provided ESA CCI LC layer assigns some of the stations to the urban area class, this is primarily due to the coarser spatial resolution of 300 m and the aerial imageries prove that the stations' neighborhood can be selected in such a way that only natural land cover (cropland, grassland) will be considered.

	KG class	ESA CCI LC (2010)	Coverage	G fraction [%]
Banloc	Dfb	urban areas	2014/11 - 2023/06	87.7
Corugea	Dfa	urban areas	2014/07 - 2021/07	83.7
Cotnari	Dfb	cropland, rainfed	2014/12 - 2023/06	80.0
Dej	Dfb	urban areas	2014/06 - 2023/06	71.6
Dumbraveni	Dfc	cropland, rainfed	2014/06 - 2021/05	75.0
Oradea	Dfb	cropland, rainfed	2014/07 - 2023/06	87.9
RosioriideVede	Dfb	cropland, rainfed	2014/06 - 2023/06	86.7
SannicolauMare	Dfa	cropland, rainfed	2015/06 - 2020/08	85.9
SatuMare	Dfb	cropland, rainfed	2014/07 - 2023/06	83.1
Tecuci	Dfb	urban areas	2014/06 - 2022/12	83.7

Table 2.1 Climate and LC classification; Cfa: humid subtropical, Dfa: humid continental (hot summer), Dfb: humid continental (warm summer)

Figure 2.2 Aerial photos, data source: *GEE*

2.2 Sentinel-1 SAR

The Sentinel-1 mission (S-1) is part of the ESA Copernicus program, designed to provide continuous all-weather, day-and-night radar imaging for land and ocean services [Sentinel-1, nd]. Launched in April 2014, S-1 consists of two identical satellites orbiting Earth, equipped with SAR instruments capable of capturing high-resolution images. These images are used for various applications, including monitoring changes in land use, detecting natural disasters, managing environmental resources, and supporting maritime surveillance. The mission contributes valuable data for scientific research, disaster management, and policy-making worldwide.

The S-1 data was gained from the C-bad Synthetic Aperture Radar Ground Range Detected (S-1 SAR GRD) image collection on the platform Google Earth Engine (GEE) for the time window 2014–2023. Of the four S-1 acquisition modes, the Interferometric Wide swath (IW) mode is mainly used for the landmasses, therefore the data was obtained

in IW mode with the high ground range resolution of 20 x 22 m [Bourbigot, 2016] and for the dual-polarization system VV and VH. The pixel size of the collection is 10 x 10 m [Bourbigot, 2016].

The GEE image collection is available already pre-processed making use of the Sentinel Application Platform SNAP S-1 Toolbox [GEE, ndc]. The pre-processing steps include the application of orbit file, GRD border noise removal, thermal noise removal, radiometric calibration, and terrain correction (orthorectification).

2.2.1 Data pre-processing

As depicted in Figure 2.2, for each point station a polygon was manually created in the GEE environment either covering the station's coordinates or is located directly next to a station in order to extract the radar signal, particularly for the neighboring fields or grassland. The S-1 SAR GRD collection was then imported filtering for the specific time range, location, acquisition mode IW and for the VV as well as VH transmitter-receiver polarisation. To reduce the speckle noise of the filtered S-1 imageries a 5x5 boxcar convolution filter was applied. The JavaScript code ready to be used in the GEE was designed and published by [Mullissa et al., 2021].

In the thesis, both, the ascending and descending orbits, were used. Since using radar data from different orbits, each scene has a different local incidence angle for the corresponding orbit and the local terrain specifications. LIA directly affects the backscatter coefficient, with greater LIA leading to lower backscatter value [Bauer-Marschallinger et al., 2021]. Hence, before exporting the S-1 SAR GRD data from GEE in order to correct the LIA impact, the information about LIA was calculated and added as the next band of each pixel for every radar scene of the collection. This is done by applying the function *addLIA.js* designed directly for the GEE JavaScript environment and developed by [Paluba, 2020], primarily dedicated to removing the normalization effect of LIA over forests. After the LIA information was added, pixel values within the polygons were sampled, once again for each image scene of the collection. This created a *FeatureCollection* which was further exported from GEE as a *csv* file containing information for the bands VV, VH and LIA. The exported dataset represents a time series of observation either early in the morning (ca. 04:30:00) or early evening (ca. 16:30:00). The temporal sampling rate varies between 6 and 12 days.

The last step in the SAR data pre-processing is the elimination of extreme backscatter coefficient values from the time series that most probably are not representa-

tive for the soil moisture conditions [Bauer-Marschallinger et al., 2019]. The SAR backscatter coefficient outliers were, therefore, eliminated from the time series following the procedure as described in studies by [Bauer-Marschallinger et al., 2019] and [Vreugdenhil et al., 2020]. In particular, a lower threshold of -20 and -26 as well as an upper threshold of -5 and -11 dB was set for the VV, VH time series, respectively.

2.3 Ancillary data

2.3.1 NDVI

To evaluate the CR variable and its suitability to describe vegetation growth and development, two ancillary datasets are used, in particular, the Normalized Difference Vegetation Index (NDVI) from Sentinel-2 (S-2) and from the Moderate Resolution Imaging Spectroradiometer (MODIS).

The S-2 NDVI product was calculated in GEE using Level-1A S-2 imageries from the harmonized GEE collection S-2 Surface Reflectance (SR) with the spatial resolution of 10 x 10 m and temporal resolution of approximately 10 days. The scenes are available from 2017/03/28. The image collection was filtered for the location of interest, for the time window 2017 - 2023 and preliminary filtered, so only up to 20% of the scene is covered by clouds. Furthermore, the cloudy pixels are filtered out by a cloud mask provided in the QA60 band of the product [GEE, nda]. The vegetation index NDVI was then computed for each scene of the collection utilizing the 4th and 8th band of the product, which represent the RED and NIR bands in the following definition of NDVI:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

The second NDVI dataset is created from the daily surface reflectance composites obtained by MODIS with a pixel size of approximately 500 m [GEE, ndb]. To preserve the underlying temporal evolution of the vegetation NDVI index, a moving average with a window size of 30 days is applied to the daily time series eliminating the noisy character.

Both NDVI datasets carry different advantages. S-2 NDVI data aligns more with the high resolution of S-1 data whereas the MODIS dataset provides higher temporal

coverage which is limited in S-2 data by lower revisit time as well as by cloudy weather. The datasets were exported from the GEE platform in the same way as the S-1 time series by extracting the mean pixel value for each polygon and further processed as a dataframe in Jupyter Notebook.

2.3.2 Vegetation Phenology Parameters

For the analysis and evaluation of the SM estimates within seasonal periods, the VPP product administrated by Copernicus Land Monitoring Service (CLMS) is deployed. The phenology data has been available since 2017. VPP product provides 13 metrics describing yearly vegetation evolution by means of Plant Phenology Index (PPI) obtained from S-2 optical data at the original 10 x 10 m resolution [Smets et al., 2023a]. Particularly, the Start of the Season (SOS) and the End of the Season (EOS) dates were extracted from the product for each of the RSMN stations. For the data extraction the Harmonized Data Access (HDA) Application Programming Interface (API) is used [Smets et al., 2023b].

2.3.3 ASCAT

The surface soil moisture product obtained by ASCAT on board of MetOp satellites and is operated by Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF), which is a part of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) [Wagner et al., 2013]. This product was used for the comparison between the original SM Change detection model without vegetation correction and for the calibrated Change detection model. The spatial resolution varies between 25 and 34 km. The product is available at a fixed grid point sampling rate of 12.5 x 12.5 km [Wagner et al., 2013] given by a WARP 5 Grid, a representation of a Discrete Global Grid [H SAF, 2022]. ASCAT SM product can be obtained from the H-SAF server in a NetCDF format. The global product is divided into many cells covering the entire Earth surface [H SAF, 2022] with one cell being defined as 5° x 5°. To access a corresponding ASCAT SM time series of a specific location, first, the unique identifier for the corresponding grid point (Grid Point Index GPI) and the cell number is found using a Python package *ascat* [Hahn et al., 2021a]. For each of the RSMN stations, a GPI and a cell number were identified and the respective NetCDF

sub-product was extracted. The time series were loaded as a *pandas* dataframe into the Jupyter Notebook environment. The time series were further filtered by a provided frost and snow mask to eliminate unreliable soil moisture measurements [Wagner et al., 2013]. The data represent measured relative SM obtained for the timestamps between 7:00-10:00 and 17:00-20:00.

2.3.4   ISMN Precipitation and Soil temperature

In addition to the SM measurements, in-situ precipitation and soil temperature data are used from the ISMN website. As with the RSMN SM database, the time series are provided at an hourly sampling rate. The precipitation dataset is used for the analysis of the model and its ability to capture the increase in SM due to rainfall. The soil temperature dataset is used in the preprocessing of ISMN SM data to eliminate SM measurements occurring during winter/early spring when soil could be frozen as those conditions cause unreliable SM estimated from satellite signal [Bauer-Marschallinger et al., 2019].

3 Methodology

This chapter details the methodology employed in the thesis, structured into four primary sections. In Section 3.1 Data Processing section is discussed 3.1.1 Local Incidence Angle normalization and 3.1.2 Merging of Time Series, ensuring comparability of SAR signal from different orbits and temporal match. Section 3.2 SM Change detection model introduces the original SM retrieval model. Section 3.3 Calibration of the SM Change detection model describes the procedures for parameter optimization to correct for the vegetation influence in the original SM change detection model. Section 3.4 Evaluation explains the criteria and methods used to assess the retrieved SM estimated from the SM Change detection model with the constant dry reference and with the calibrated dry reference.

3.1 Data Processing

3.1.1 Local Incidence Angle normalization

LIA is an important step in the processing of SAR data. It enhances the comparability of the data obtained from different orbits. Normalizing LIA reduces the influence of angle variations, facilitating accurate comparison and integration of data. The linear relation method to normalize LIA similarly to [Bauer-Marschallinger et al., 2021] was applied (Fig. 3.1a,b).
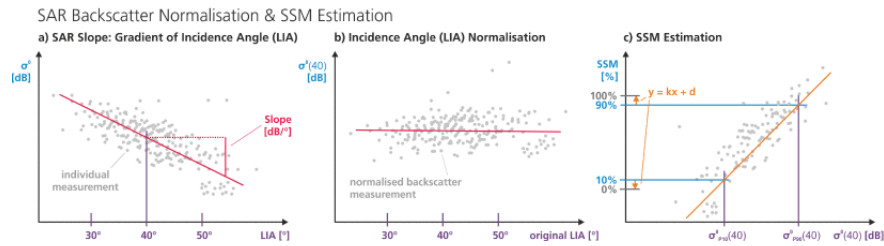
Figure 3.1   LIA normalization (a,b); estimation of dry and wet reference (c), source: [Bauer-Marschallinger et al., 2019]

First, the slope parameter of β was determined using Simple Linear Regression for LIA (x) and SAR backscatter coefficient (y). Python package function *LinearRegression* from the *sklearn.linear_model* module was used[1]. The equation 3.1 introduced by

---

[1]`https://scikit-learn.org/0.20/modules/generated/sklearn.linear_model.`
`LinearRegression.html`

[Bauer-Marschallinger et al., 2021] was then used to normalize the SAR signal.

$$\sigma^0_{\theta_{ref},t} = \sigma^0_{\theta,t} - \beta * (\theta - \theta_{ref}), \tag{3.1}$$

where $\sigma^0_{\theta_{ref},t}$ is the normalized backscatter coefficient, $\sigma^0_{\theta,t}$ is the backscatter coefficient for each scene and with a respective LIA (θ), β is the estimated slope parameter from the linear regression, θ is the respective LIA and θ_{ref} is the reference angle defined as a rounded mean of LIA value range. The correction was performed for both polarizations and yields $\sigma^0_{VV_{norm}}$ and $\sigma^0_{VH_{norm}}$.

After LIA correction, CR was calculated using the following definition

$$CR = \sigma^0_{VH_{norm}} - \sigma^0_{VV_{norm}} [dB]. \tag{3.2}$$

3.1.2 Merging of Time Series

Having all time series pre-processed, the data can be merged. ISMN SM, precipitation and soil temperature were merged first since they are all provided at the hourly sampling rates. Satellite data originates from a coarser temporal resolution (ca. 1 - 12 days). Thus, the closest ISMN measurement was found for each S-1 data point. To guarantee that the satellite captures the current soil moisture conditions the closest measurement in the past (backward direction) was found. In the same way, ASCAT SM time series were matched with the dataframe. As mentioned above, S-1 satellites fly over the areas of interest around 4:00 in the morning and 16:00 in the early evening. ASCAT sensor dates the measurements over the areas between 7:00–10:00 and 17:00-20:00. Therefore, the closest data point from the ASCAT SM dataset was matched with the dataframe. The remaining datasets (S-2 NDVI, MODIS NDVI) were merged with a left join without interpolating for the missing values. Data without ISMN SM measurements or with ISMN soil temperature $< 4°C$ were eliminated from the dataset [Hahn et al., 2021b].

3.2 SM Change detection model

The original Change detection model was developed by [Wagner et al., 1999], initially to retrieve SM from the ERS scatterometer. The model is based on the assumption that the backscatter coefficient can be linearly related to the dielectric properties of soil which are in turn related to the water content in soil [Wagner et al., 2013]. The backscatter coefficient is scaled between so called wet ($\sigma^0_{norm,wet}$) and dry reference ($\sigma^0_{norm,dry}$). They

represent the long-term driest and wettest soil conditions and should be retrieved from a multi-year (>10 years) time series [Wagner et al., 2013]. Following the procedure by [Bauer-Marschallinger et al., 2019] (see Fig. 3.1c) the dry and wet references are calculated as the 10th and 90th percentile of the backscatter time series. The S-1 data archive is approaching the 10-year threshold, but as [Bauer-Marschallinger et al., 2019] states the determination of the dry and wet reference may still be limited due to the insufficiently long archive and possible outliers due to extreme situations (e.g., frozen soil) that could be misinterpreted as referred reference. The [Bauer-Marschallinger et al., 2019] procedure using the 10th and 90th percentiles seems to reduce similar influences. The percentile backscatter coefficient is further linearly extrapolated to the 0th and 100th percentile applying the following equations from [Bauer-Marschallinger et al., 2019]

$$k = \frac{90\% - 10\%}{\sigma^0_{90,norm} - \sigma^0_{10,norm}}, \tag{3.3}$$

$$d = 90\% - k * \sigma^0_{90,norm}, \tag{3.4}$$

$$\sigma_{norm,dry} = \frac{0\% - d}{k} \qquad [dB], \tag{3.5}$$

$$\sigma_{norm,wet} = \frac{100\% - d}{k} \qquad [dB]. \tag{3.6}$$

The defined $\sigma_{dry,norm}$ and $\sigma_{wet,norm}$ can later be inserted in the SM Change detection equation:

$$SM_t = \frac{\sigma^0_{norm,t} - \sigma^0_{norm,dry}}{\sigma^0_{norm,wet} - \sigma^0_{norm,dry}} \qquad [-]. \tag{3.7}$$

Since the dry and wet reference is defined as 10th and 90th percentile, it can happen that the backscatter coefficient will exceed the scaling value range calculated as described. In this case, the computed soil moisture estimates are set to 0 or 1, respectively.

The computed soil moisture is only relative to the value range defined by the wet and dry reference and is expressed in relative number 0–1. To compare the measured in-situ SM (Θ_v) given in the volumetric water content unit $[m^3.m^{-3}]$ and unitless relative soil moisture estimates, both variables must be described in the same units. There are two possible ways to achieve this. The first is to multiply the relative soil

moisture by soil porosity thus gaining the absolute soil moisture. The soil porosity dataset used for this conversion is the HWSD (version 1.1 with 1 km resolution) [Dorigo et al., 2021]. However, the coarse resolution of 1 km can introduce another source of error by not characterizing the actual local porosity. The next option is to standardize the retrieved relative SM values to have the same distribution as the measured in-situ data [Wagner et al., 2019]. This fitting method is described as follows

$$\Theta_{v,t} = \frac{SM_t - \overline{SM}}{\sigma_{SM}} * \sigma_{\Theta_v} + \overline{\Theta_v}, \tag{3.8}$$

where $\Theta_{v,t}$ is the observed soil moisture expressed in volumetric units, SM_t is the relative SM retrieved from the satellite, $\overline{SM}$ and $\overline{\Theta_v}$ are the means of the retrieved SM and the observed in-situ SM, σ_{SM} and σ_{Θ_v} are the variances of the retrieved relative SM and the in-situ SM. The latter option is implemented in the thesis as it is not dependent on ancillary dataset.

3.3   Calibration of SM Change detection model

The main aim of the thesis is to calibrate the original Change detection model using the CR variable to modify the constant dry reference. Several studies proved the correlation between CR and vegetation development by comparing the radar variable with different vegetation indices - NDVI [Veloso et al., 2017], VOD [Vreugdenhil et al., 2020], Enhanced Vegetation Index (EVI) [Ma et al., 2024]. The new calibrated dry reference should change with time. Particularly, an increase of the new calibrated reference is expected with vegetation growth as well as a decrease during periods without vegetation. The proposed dry reference $\sigma^0_{norm,dry-cal}$ can be expressed mathematically as follows

$$\sigma^0_{norm,dry-cal} = a * \overline{CR}_{31} \tag{3.9}$$

The parameter a is determined by an optimization process, explained in more detail in next section 3.3.1.

The proposed variable $\overline{CR}_{31}$ is defined in two steps. First, the value range of CR needs to be shifted vertically along the y-axis (standing for σ^0_t in [dB]). This step is necessary in order to preserve the low backscatter value of dry reference. The mean of the shifted CR time series equals now the defined 0th percentile of the backscatter

coefficient (Fig. 3.2).

Afterward, the $\overline{CR}_{31}$ is defined as a 31-days moving average in order to smooth out short-term fluctuations and highlight underlying longer-term trends in the shifted CR time series.

Figure 3.2 below displays the described steps before the optimization - the initial CR time series and constant dry reference without pre-processing (a), the vertical shift (b), 31-days moving average smoothing and an example of a calibrated dry reference (c).
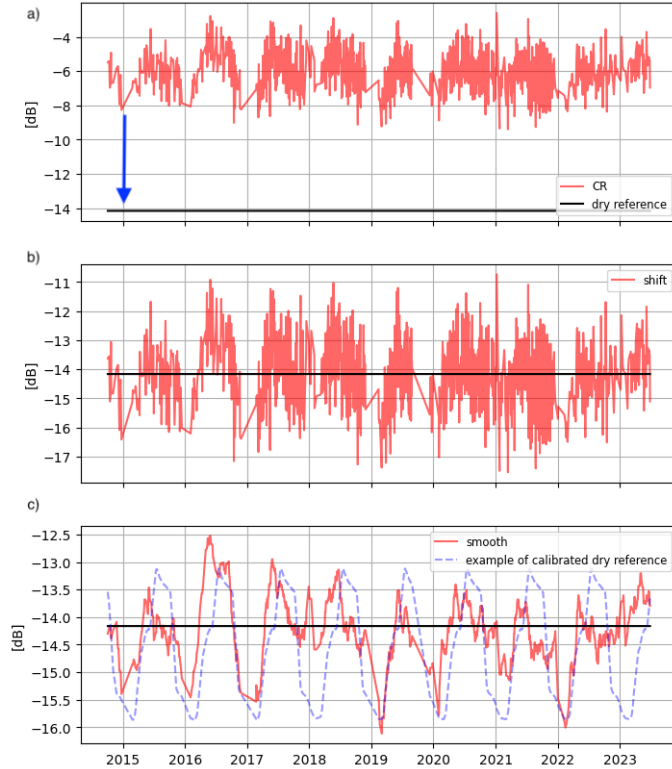
Figure 3.2 Shifting and smoothing of CR time series

The variable $\sigma^0_{norm,dry-cal}$ from Eqaution 3.9 serves as a replacement for the original time-invariant $\sigma^0_{norm,dry}$ in the Change detection model (see also Equation 3.7)

$$SM_t = \frac{\sigma^0_{norm,t} - \sigma^0_{norm,dry-cal}}{\sigma^0_{norm,wet} - \sigma^0_{norm,dry-cal}}. \qquad [-] \qquad (3.10)$$

Similarly, the relative values can be normalized via the Equation 3.8 to the absolute units of $m^3.m^{-3}$.

3.3.1 Optimization process

The calibration was achieved using the optimization method Limited-memory Broyden-Fletcher-Goldfarb-Shanno *L-BFGS-B*. The function optimized by this algorithm is $RMSD$ (see Section 3.4.1) between the observed in-situ SM and the retrieved SM.

The optimization method *L-BFGS-B* is implemented in the in-built function *minimize* from the Python library *SciPy.optimize* [2]. The function requires the following arguments - *fun, x0, args, method*. The *fun* is the objective function to be minimized ($RMSD$). The optimization algorithm attempts to find the optimal parameter values in order to minimize the objective function.

The next argument *x0* stands for the initial *a* parameter guess. The multiplicative parameter *a* is optimized for each station. Moreover, the parameter *a* is optimized for each day of the year (1-365). The optimized parameter changes for each observation and can be defined as

$$a_N = \begin{pmatrix} a_{t_1} \\ a_{t_2} \\ \vdots \\ a_{t_N} \end{pmatrix},$$

where N is the length of the time series. As the initial guess for the *a* parameter is an *N*-dimensional vector of ones

$$1_N = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

The *args* argument is a set of optional arguments required for the computation of the objective function *fun*. For the presented case, the variables necessary for the model (*VV_norm*, *CR*, the original constant $\sigma_{norm,dry}$, $\sigma_{norm,wet}$) are passed to the *args* argument.

---

[2]`https://docs.scipy.org/doc/scipy/reference/optimize.minimize-lbfgsb.html#`
`optimize-minimize-lbfgsb`

3.4 Evaluation

For the quantitative evaluation of models' performances, two metrics are computed namely $RMSD$ and Pearson correlation coefficient r.

3.4.1 RMSD

The root mean square deviation $RMSD$ is a statistical metric used to measure the mean absolute deviation between values predicted by a model and the values observed. It provides a single value that represents the mean magnitude of these differences. It is expressed as

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - Y_i)^2}.$$
(3.11)

X_i is the observed reference variable, Y_i the retrieved estimate and N stands for the number of data samples. Common target $RMSD$ value is 0.004 $m^3.m-3$ [Gruber et al., 2020].

3.4.2 Pearson correlation coefficient r

The Pearson correlation coefficient (r) is a statistical measure that assesses the strength and direction of the linear relationship between two variables. It ranges from -1 to 1, where 1 indicates a perfect positive linear correlation, -1 indicates a perfect negative linear correlation and 0 signifies no linear correlation. The coefficient is defined as

$$r = \frac{\sum_{i=1}^{N}(X_i - \overline{X}) * (Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \overline{X})^2} * \sqrt{\sum_{i=1}^{N}(Y_i - \overline{Y})^2}},$$
(3.12)

where X_i is the observed reference variable, Y_i the retrieved estimate and $\overline{X}$ and $\overline{Y}$ denote the respective mean variable value.

4 Results

The following chapter is devoted to describing the results of the thesis experiment. Using the S-1 SAR data, soil moisture was estimated by applying the SM Change detection model (Section 4.1). A new calibrated dry reference $\sigma^0_{norm,dry-cal}$ was determined using optimization of the parameter a and subsequently used in the SM Change detection model instead of the original constant dry reference (Section 4.2). The horizontal line in Tables 4.1 and 4.2 outlines the boundary between stations meeting the required target threshold of 0.04 $m^3.m^{-3}$ (Section 3.4.1).

The main problem addressed in the thesis is the correction of vegetation influence on soil moisture determination from satellite data and its elimination. Therefore, the results for the original SM Change detection model as well as the calibrated Change detection model are compared with each other for both seasonal and off-seasonal periods. In this way, it is possible to evaluate the hypothesis whether the calibration has a more significant benefit to the model during the presence and growth of vegetation (Section 4.3).

In addition, the results of both models were compared with the established ASCAT SM product at 12.5 km resolution (Section 4.4). Again, the calibration contribution compared to the original SM Change detection and to the ASCAT SM product is evaluated for the whole dataset as well as for the season and off-season.

4.1 Results for SM Change detection model

SM Change detection model with the constant dry reference retrieved soil moisture with $RMSD$ between 0.021 and 0.091 $m^3.m^{-3}$ (Table 4.1). Only two out of the total 10 stations reached $RMSD$ smaller than 0.04 $m^3.m^{-3}$ being the target threshold. Station *RosioriideVede* achieves the lowest $RMSD$ with a relatively high Pearson correlation of 58%. The *Corugea* station yields low $RMSD$ of 0.034 $m^3.m^{-3}$ with moderate correlation of 41%. *Cotnari* station, albeit achieving low $RMSD$ of 0.041 $m^3.m^{-3}$, shows a very low correlation of less than 9%. Stations *Tecuci*, *Oradea* and *Dumbraveni* obtain all the same $RMSD$ of 0.047 $m^3.m^{-3}$ with different correlation coefficients. The highest overall correlation between the satellite estimates and ground observations of almost 60% is reached for the *Oradea* station. Station *SannicolauMare* exhibits similar $RMSD$ (0.055 $m^3.m^{-3}$) and correlation coefficient (28.6%) value to the overall mean $RMSD$ of 0.053 $m^3.m^{-3}$ and mean correlation of 33%. Stations *Banloc*, *SatuMare* and especially

station *Dej* obtain the highest *RMSD* and weak correlation up to 32%.

Station	RMSD [$m^3.m^{-3}$]	r [%]
RosioriideVede	0.021	58.12
Corugea	0.034	41.43
Cotnari	0.041	8.89
Tecuci	0.047	30.08
Oradea	0.047	59.3
Dumbraveni	0.047	23.81
SannicolauMare	0.055	28.6
Banloc	0.072	24.61
SatuMare	0.076	31.56
Dej	0.091	32.18
Mean	**0.053**	**33.86**

Table 4.1 Results - SM Change detection model with constant dry reference

4.2 Results for the calibrated SM Change detection model

The calibrated SM Change detection model with the new dry reference decreased the mean *RMSD* by 7.5% (0.004 $m^3.m^{-3}$) relative to the original SM model and increased the correlation coefficient by approximately 8.7% (Table 4.2). However, the calibration leads to a decrease in *RMSD* under the threshold of 0.04 $m^3.m^{-3}$ only for the station *Cotnari* compared to the original model.

The highest percentual decrease in *RMSD* is achieved for the stations *Banloc* (by 13.89%) and *Cotnari* (by 9.76%) and in absolute units again for the station *Banloc* (from 0.072 $m^3.m^{-3}$ to 0.062 $m^3.m^{-3}$), *Dej* (from 0.091 $m^3.m^{-3}$ to 0.084 $m^3.m^{-3}$) and *SatuMare* (from 0.076 $m^3.m^{-3}$ to 0.07 $m^3.m^{-3}$). The original model performs poorly for the stations *Banloc*, *Dej* and *SatuMare* (Table 4.1). The calibration was able to enhance SM retrieval the most for these stations. The correlation coefficient also improved especially for these stations by almost 20%, 10% and more than 9%, respectively. For the stations *RosioriideVede* and *Corugea*, neither *RMSD* nor correlation improves distinctly after the model calibration compared to the original model as the original model performed already very well for these locations. Great improvement in terms

of the correlation coefficient is observed for the station *Cotnari* from ca. 8% to more than 25%. This station can be specifically distinguished from all the other stations by being located next to a vineyard (Fig. 2.2). The remaining stations *Oradea*, *Tecuci*, *Dumbraveni* and *SannicolauMare* show a slight decrease of $RMSD$ and an increase in correlation between 5.4 and 7.9%.

Generally, model calibration leads to better performance of the SM Change detection model, though the highest improvement is observed for the stations where the original model performed poorly. On the other hand stations with relatively good results yielded from the original model are enhanced by the calibration only a little or moderately.

Station	RMSD	Δ RMSD [%]	r	Δ r [%]
RosioriideVede	0.02	-4.76	61.13	+3.01
Corugea	0.032	-5.88	45.27	+3.84
Cotnari	0.037	-9.76	25.48	+16.59
Oradea	0.044	-6.38	64.72	+5.42
Tecuci	0.045	-4.25	35.75	+5.67
Dumbraveni	0.045	-4.26	29.77	+5.96
SannicolauMare	0.052	-5.45	36.51	+7.91
Banloc	0.062	-13.89	44.42	+19.81
SatuMare	0.07	-7.89	41.69	+10.13
Dej	0.084	-7.69	41.63	+9.45
Mean	**0.049**	**-7.53**	**42.64**	**+8.78**

Table 4.2   Results - calibrated SM Change detection model with modified dry reference

4.3   Temporal evaluation

The model calibration was applied to the entire time series. However, the main impact of the calibration should be detected over the seasons during vegetation growth. Hence, the evaluation metrics are computed separately for season (Table 4.3-4.4) and off-season periods (Table 4.5 - 4.6) to assess the temporal impact of the calibration. The VPP dataset was used to identify seasons for the locations. The VPP dataset is available only for the years 2017–2022 and therefore this evaluation is available for this specific time window. Both winter and summer seasons are considered. Table 4.3-4.6

are ordered by the highest enhancement caused by calibration (Δ column).

After selecting seasons and off-seasons, it is clear that for the majority of stations the original SM Change detection model has more complications with the off-season periods. Mean $RMSD_\text{season}$ is 0.048 $m^3.m^{-3}$ and r_season is 38.95 % compared to the mean $RMSD_\text{off-season}$ of 0.057 $m^3.m^{-3}$ and $r_\text{off-season}$ of 29.74%. By calibrating the model, it was possible to reduce the imbalance between the two periods and to achieve comparable $RMSD$ ($RMSD_\text{season} = 0.045 m^3.m^{-3}$, $RMSD_\text{off-season} = 0.057 m^3.m^{-3}$) and correlation values ($r_\text{season} = 45.28\%$, $r_\text{off-season} = 41.76\%$). In general, the calibration appears to have identified the periods that have the greatest impact on the poor performance of the original model and these have been adjusted the most.

As is the case for the entire time series, the analysis of the individual locations shows large differences between stations. The results for single stations are described in the next Sections 4.3.1 and 4.3.2.

4.3.1 Season

Calibration of the model results in an average decrease of $RMSD_\text{season}$ by 5.64% (0.003 $m^3.m^{-3}$) compared to the original SM Change detection model (Table 4.3) and to an average increase in r_season by 6.33% (Table 4.4). The greatest improvement in terms of $RMSD$ is observed for the station *Corugea* ($\Delta RMSD_\text{season} = -9.68\%$) but correlation is not improved a lot by calibration ($\Delta r_\text{season} = +3.89\%$). Station *Corugea* is burdened by many ISMN data dropouts and thus the entire time series is very short (257 data points) compared to the remaining stations. Selecting only seasonal time windows reduces the number of data and therefore puts an even greater burden on the interpretation of the results. A similar limitation occurs for *Dej* (460 data points), *SannicolauMare* (561 data points) and *Dumbraveni* (684 data points) stations, although the length of the total time series is not as low as for *Corguea* station. All of the mentioned stations appear at the bottom of Table 4.3 exhibiting no or only a low decrease in $RMSD_\text{season}$ between 0 to 4.6%. The remaining time series consists of more than 1 000 data.

The second highest decrease in terms of $RMSD$ is observed for *Cotnari* station with a distinct vineyard vegetation by more than 8% for $RMSD_\text{season}$ and by the greatest increase in r_season by more than 12%. Interestingly, the original model performs distinctly better during the seasonal periods than outside the season ($r_\text{season} = 21.97\%$

vs. $r_{\text{off-season}} = -10.7\%$).

The remaining stations exhibit a decrease in $RMSD_{\text{season}}$ between 2.3% up to 8% and an increase in correlation coefficient between 2% to 9% with the lowest improvement particularly for the stations with short time series.

Station	Original	Calibrated	Δ RMSD [%]
Corugea	0.031	0.028	-9.68
Cotnari	0.035	0.032	-8.57
SatuMare	0.063	0.058	-7.94
Tecuci	0.052	0.048	-7.69
Banloc	0.054	0.05	-7.41
RosioriideVede	0.019	0.018	-5.26
Dej	0.087	0.083	-4.6
Oradea	0.046	0.044	-4.35
Dumbraveni	0.042	0.041	-2.38
SannicolauMare	0.05	0.05	0.0
Mean	**0.048**	**0.045**	**-5.64**

Table 4.3 Seasonal $RMSD$ for original and calibrated model

Station	Original	Calibrated	Δr [%]
Cotnari	21.97	34.51	+12.54
Tecuci	19.49	28.85	+9.36
SatuMare	33.54	42.81	+9.27
Banloc	37.43	45.05	+7.62
Dej	29.96	35.97	+6.01
Oradea	53.85	58.57	+4.72
Dumbraveni	28.72	33.31	+4.59
Corugea	60.93	64.82	+3.89
RosioriideVede	57.85	61.11	+3.26
SannicolauMare	45.78	47.82	+2.04
Mean	**38.95**	**45.28**	**+6.33**

Table 4.4 Seasonal r for original and calibrated model

43

4.3.2 Off-season

Mean $RMSD_\text{off-season}$ decreased by more than 9% (0.05 $m^3.m^{-3}$) in comparison to the original model (Table 4.5) and the calibration improved by 12% (Table 4.6). The biggest improvement in off-season results occurred at *Banloc* station. This station also showed a slightly above average improvement during the season (Tables 4.3,4.4). From the results in Table 4.2, it can be seen that the calibration had the greatest impact on this station and produced the highest change in $RMSD$ and correlation. Therefore, the calibration for this station had an impact for both time periods, although here too the effect is more visible for the off-season. It should also be noted that both the seasonal and off-seasonal time series are almost identically long and thus there is likely no bias in the results due to the unbalanced dataset size. A similar improvement for the season and off-season can be seen for the *Cotnari* station.

Stations *Sannicolau* and *Dej*, which are also at the top of Tables 4.5 and 4.6, show only average to below-average improvement using calibration during saeson (Table 4.3,4.4). Nevertheless, the calibration led to overall high changes across the time series (Table 4.2) and therefore it is evident that the largest change occurred in the off-season.

In general, model calibration was most helpful in the off-season as defined by the VPP dataset. The only site for which the opposite behavior, and thus a greater improvement during the season than the off-season, is observed is the *Tecuci* station with a greater $\Delta RMSD_\text{season}$ of more than 7% and Δr_season of more than 9% (Table 4.3, 4.4) compared to only a marginal $\Delta RMSD_\text{off-season}$ of 2% and $\Delta r_\text{off-season}$ of less than 3%.

Station	Original	Calibrated	Δ RMSD [%]
Banloc	0.087	0.071	-18.39
SannicolauMare	0.054	0.047	-12.96
Cotnari	0.049	0.043	-12.24
Dej	0.088	0.079	-10.23
Oradea	0.047	0.043	-8.51
SatuMare	0.092	0.085	-7.61
RosioriideVede	0.023	0.022	-4.35
Dumbraveni	0.05	0.048	-4.0
Corugea	0.034	0.033	-2.94
Tecuci	0.042	0.041	-2.38
Mean	**0.057**	**0.051**	**-9.54**

Table 4.5 Off-seasonal $RMSD$ for original and calibrated model

Station	Original	Calibrated	Δr [%]
Banloc	16.55	44.23	27.68
Cotnari	-10.7	11.73	22.43
SannicolauMare	16.31	37.62	21.31
Dej	32.66	44.15	11.49
SatuMare	29.68	40.5	10.82
Dumbraveni	21.16	29.46	8.3
Oradea	62.94	69.24	6.3
Corugea	30.43	35.78	5.35
RosioriideVede	57.37	60.89	3.52
Tecuci	41.04	43.95	2.91
Mean	**29.74**	**41.76**	**+12.01**

Table 4.6 Off-seasonal r for original and calibrated model

4.4   Cross-Comparison

The results in the presented work from high spatial resolution S-1 data are compared
in this section with the well-established 12.5 km resolution ASCAT SM product for the

entire time series, season and off-season (Fig. 4.1).

Overall, it can be assessed that the calibrated model compared to the original model led to a slight decrease in $RMSD$ and a more pronounced increase in correlation. Despite the overall improvement of results through calibration as well as achieving a higher overall correlation r at 2 stations (Fig. 4.1d), the ASCAT SM product shows a slightly lower mean and median $RMSD$ as well as higher correlation r in general. For the season, the mean and median $RMSD$ values (Fig. 4.1b) as well as the correlation (Fig. 4.1e) between the calibrated model and ASCAT SM are comparable. Nevertheless, the differences between stations are somewhat smaller in the case of ASCAT SM than between the results from the calibrated S-1 model. Regarding the off-season results, it is interesting to observe that all three models lead to large differences in $RMSD_{\text{off-season}}$ between stations which is reflected in the width of the boxplot (Fig. 4.1c). ASCAT SM, as well as the results of both S-1 models, show a higher $RMSD_{\text{off-season}}$. Nevertheless, $r_{\text{off-season}}$ for ASCAT SM is much higher than $r_{\text{off-season}}$ of S-1 model even after calibration (Fig. 4.1f). The lowest $r_{\text{off-season}}$ for ASCAT SM is more than 35% (*Corugea*) which is still higher than the result for the same station after calibration (30%, see Table 4.6) and much more than the lowest $r_{\text{off-season}}$ for the S-1 calibrated model (*Cotnari*, 11%).
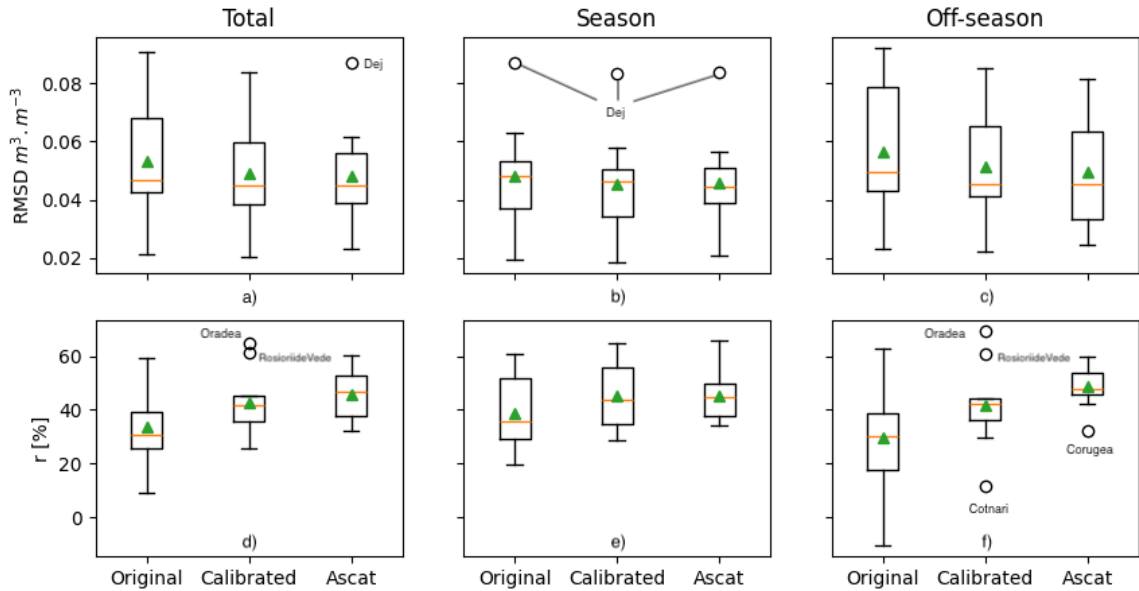

Figure 4.1 Cross-comparison of all three products with displayed mean (green triangle) and median (orange line)

5 Discussion

The presented thesis studies the potential of CR to be used for vegetation correction in soil moisture retrieval at field-scale by C-SAR S-1. This was done by modifying the constant dry reference by the smoothed and shifted annual monthly CR time series. The new time-varying dry reference was then multiplied by a calibrated parameter time series to optimize its amplitude. The results are evaluated in the temporal domain, comparing the in-situ and satellite SM retrieval within a long-term time span (>8 years, if in-situ SM is available). Spatial agreement between the in-situ data and satellite estimates is not included because for each station only one point-like measurement is available.

5.1 SM Change detection model

The S-1 Change detection model with static dry reference achieved the worst results among the three models. The constant dry reference fails to capture changes in the land surface caused by vegetation that varies over time. Thus, a bias element is introduced into the model that causes either under- or overestimation of SM estimates. As mentioned in the Results section 4 there are large differences between stations and for the cause of this, it is necessary to evaluate the time series for each site individually.

The $RMSD$ for the Dej station is significantly higher than for the remaining stations. This station is heavily burdened by the low number of data points in the time series due to the frequent absence of in-situ SM measurements. Upon closer analysis of the time series, strong noisy behavior of the SAR time series can be observed. The noise causes extreme fluctuations in the backscatter signal (Fig. 5.1). This is not a positive or negative bias, but a chaotic fluctuation. It is this behavior between successive data that could indicate insufficient LIA correction between relative orbits. The LIA effect has not been removed completely and it is likely that particularly for this station it causes problems. The noise is particularly pronounced during the decline of the in-situ SM measurements (dry periods) and their constantly low value (highlighted in black). The SAR signal is complex and reflects not only the volume of water in the soil but also changes in the surface, its roughness or the structure of the vegetation. Local soil properties can also have an impact. The fluctuations may also be a reflection of these local factors appearing during drier conditions. The noise is observable for the original model as well as the calibrated one, while it becomes slightly more pronounced

after calibration. This is a consequence of the narrowing of the so-called *sensitivity* or *dynamic range* which represents the range between dry and wet reference (Fig. 5.1c). As explained by [Pfeil et al., 2018] if the sensitivity is too narrow it can result in high noise in the SM retrieval.
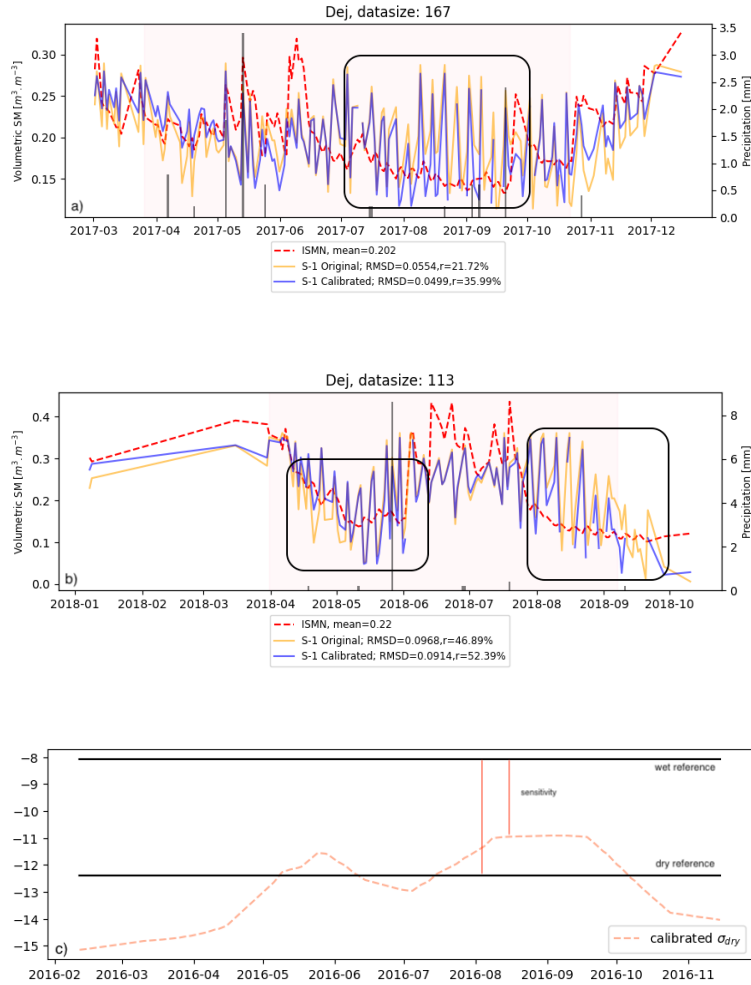

Figure 5.1 Noise presence in SM satellite estimates under dry conditions, (a) 2017, (b) 2018, (c) static and calibrated dry reference with highlighted change in the *sensitivity* (red vertical line)

At the *Cotnari* station, the opposite result can be observed, namely a relatively low *RMSD* but at the same time the lowest correlation. Low correlation between satellite time series and in-situ measurements was reported also in the study by [Bauer-Marschallinger et al., 2019] with the lowest correlation of 11%, similar to the correlation coefficient of *Cotnari* station. Even though the correlation improved the result considerably, the location represents an area where the backscatter coefficient does not appear to capture the in-situ SM temporal course. *Cotnari* station is characterized

by vineyards. Vineyards have a characteristic field structure defined by inter-row spacing or their orientation and jumps in between them as well as the vegetation characteristic structure with vertical stems and branches and woody plant components. This location is therefore particularly challenging due to the potential roughness aspect as well as vegetation attenuation.

5.2   Calibration and vegetation correction

The results suggest potential in the here introduced vegetation correction, albeit with some limitations and with differences between single locations indicating the relevance of the local properties. Generally, the calibration led to a higher correlation coefficient by more than 8% and to lower $RMSD$ by more than 7%. Also, model calibration benefited especially in locations where the original Change detection model fails. Compared to the vegetation correction in ASCAT SM retrieval by [Pfeil et al., 2018], the increase in correlation due to vegetation correction is smaller (9% vs. 15%). [Bhogapurapu et al., 2022] performed vegetation correction for SM retrieval using the Change detection model and Dual-pol Radar Vegetation Index (DpRVIc) for the correction. They compared this methodology with the correction exploiting the NDVI from optical data and achieved better accuracy for the DpRVIc with very high correlation of more than 80%. Their results point to the suitability of SAR variables for monitoring and eliminating the impact of vegetation. Such high correlation coefficients were not obtained in the current thesis, even though the spatial resolution between the current thesis and the mentioned study is more comparable than that in the study by [Pfeil et al., 2018]. The study area (included in the Texas Soil Observation Network TxSON) in the article by [Bhogapurapu et al., 2022] is used as calibration and validation site for the SMAP SM product. Locations used in the study are distributed over much larger area of the entire Romania and are affected by local and regional phenomena (e.g., weather, vegetation type, local soil properties) and therefore the applied S-1 model is probably less suited for some of the locations than for the others.

The station *Banloc* appears overall to benefit the most from the correction. From the Google Earth Pro aerial imageries, the area represents grassland. Figure 5.2a displays the 2017 time series for both models and the observed in-situ SM. Interesting is the ability of the calibration to eliminate the effect of the underestimation during spring months (red rectangle). The presence of vegetation leads in general to either over- or underestimation. The vegetation water contect might increase the backscatter coefficient

and therefore falsely be interpreted as the soil moisture content. This would lead to an overestimation of the SM estimates. The underestimation effect was explained by [Pfeil et al., 2018] as a consequence of the vegetation structural changes from grass-like to vertical which lead probably to a decrease in received backscatter signal in the SM retrieval. These effects were studied as well by [Veloso et al., 2017] and the CR variable appeared very well suited for vegetation monitoring, particularly for this vegetation growth stage. [Pfeil et al., 2018] performed a correction based on a cross-angle principle for the ASCAT SM product and detected a similar pattern of SM underestimation during the spring times along with overestimation of microwave VOD of vegetation for that time. Using CR as the dry reference as presented in the current thesis successfully decreased this discrepancy bias by enlarging the sensitivity between the dry and wet reference (Fig. 5.2b).

As for the overestimation of SM estimates (black rectangle), here too the result was improved with the help of calibration. However, it is very necessary to correctly set the maximum range of the parameter during calibration so that the calibrated dry reference is not too high. This would cause either the mentioned increased noise due to the small range between the references or even impossible relative SM estimates above 100 %.
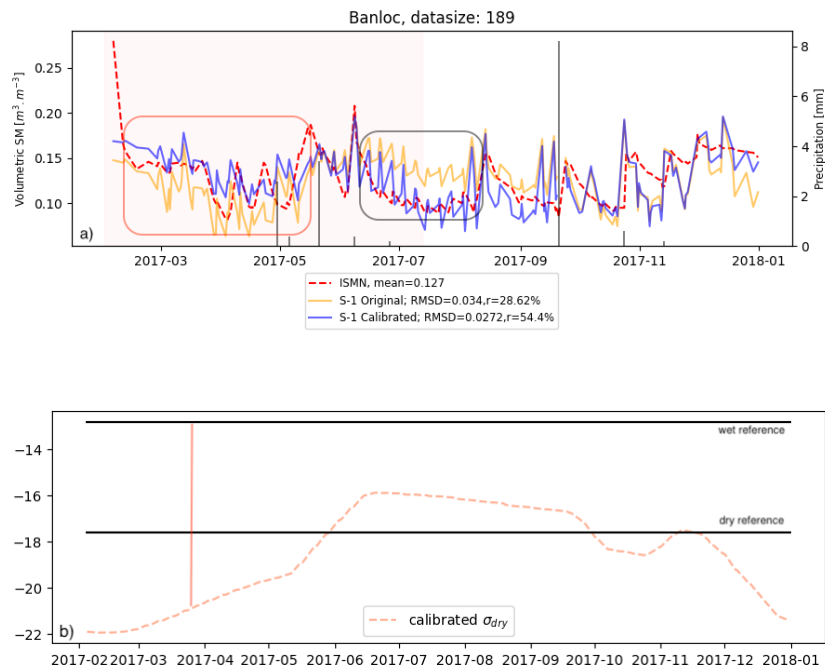

Figure 5.2   (a) SM time series for *Banloc* station (2017), (b) static and calibrated dry reference

The assumption of the proposed correction is the potential of CR to describe

vegetation growth. SAR CR proved to be correlated to MODIS NDVI either moderately or even strongly (Fig. 5.3) and the correlation coefficient reached high values. Similar results for the correlation gained [Veloso et al., 2017] when comparing seasonal time series of different winter and summer crop types (wheat, rapeseed, barley, maize, soybeans, sunflower). Correlations between 63% and almost 91% that are comparable with the correlation by Veloso et al. (74-89%) were found. However, they reported differences between single crop types. For sunflower fields correlation between CR and NDVI was almost non-existent (0.08%) indicating that the vegetation type has an immense impact on the ability of CR to monitor vegetation or to be used for vegetation correction. CR should correlate well with the vegetation growth if the VV backscatter coefficient component reflects mostly the signal from the ground and the VH component the vegetation volume scattering. Then, CR obtains higher values (Section 3.2, $\sigma^0_{VH_{norm}} - \sigma^0_{VV_{norm}}$). In the case of sunflower, the VV backscatter coefficient stems mostly from the volume scattering as well as VH backscatter and, hence, CR cannot suppress the ground effect and highlight the vegetation.

The correlation with the NDVI from S-2, though stemming from sensors with more comparable spatial resolution, is in general lower than for NDVI MODIS (Fig. 5.4). The main problem in the comparison between NDVI from S-2 and CR S-1 si the small number of samples for NDVI due to cloudy weather conditions. The *Corugea* station even shows a negative correlation between CR and NDVI, which is most likely due to the low number of both S-2 NDVI data and ISMN data. Nevertheless, the insight into the comparison of CR with another NDVI dataset, at a superior spatial resolution, reveals that low correlation coefficients are found for station *RosioriideVede* and *Dumbraveni*. Both localities do not benefit from the calibration significantly as shown in the Section 4. However, it is very likely not a consequence of the weak correlation between S-2 NDVI and S-1 CR, but other unexplained reasons, since MODIS NDVI points to a high correlation on the contrary.
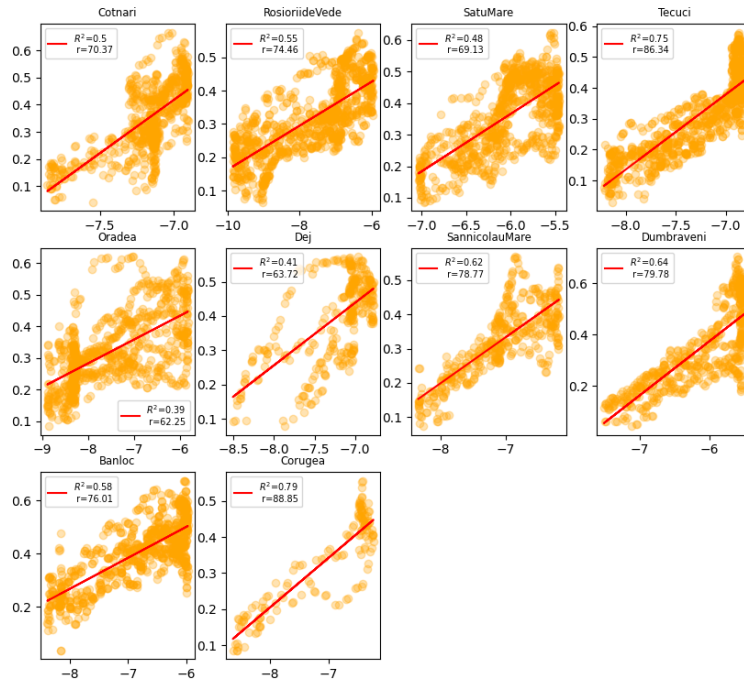
Figure 5.3 Correlation coefficient r and R^2 between NDVI MODIS and S-1 CR
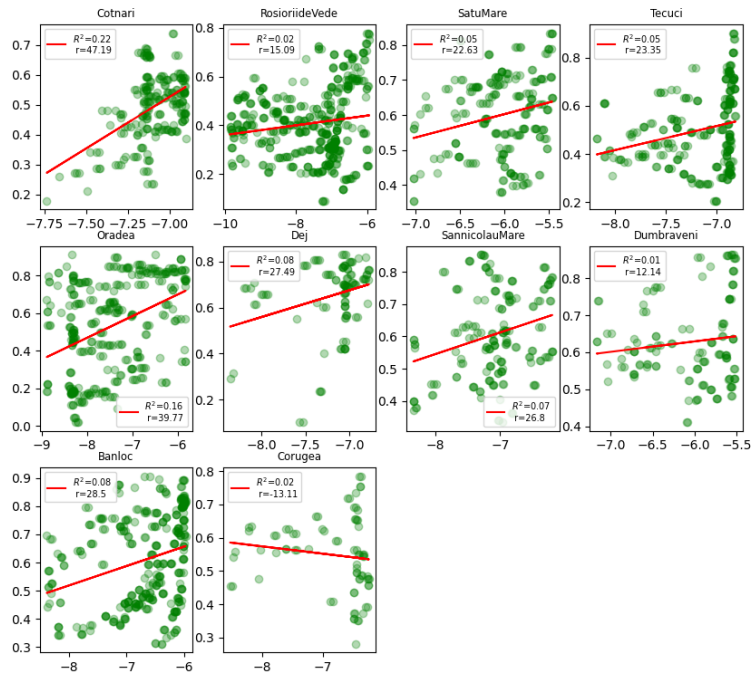

Figure 5.4 Correlation coefficient r and R^2 between S-2 NDVI and S-1 CR

The relationship between the SAR CR and the optical dataset is also relevant considering the use of the VPP dataset, which is created from optical data. The initial assumption that calibration would help most during the season defined by VPP was not confirmed. This can be largely caused by the different sources of both datasets. As

stated by [Veloso et al., 2017], CR correlates very well with NDVI, but their time course is not completely identical. NDVI is characterized by a rapid increase already at the initial plant appearance, while CR is influenced by structural changes in the vegetation in the initial phase of growth. Also, after harvest, NDVI decreases immediately due to changes in absorption and reflectance of NIR and RED, while CR can still be affected either by the presence of photosynthetically inactive vegetation residues or by changes in roughness. Although these factors appear outside the season of vegetation growth, they can still affect SM retrieval and therefore their correction is beneficial.

5.3 Cross-comparison

From the comparison of the work results with the ASCAT SM product, it is obvious that the ASCAT SM product leads to the best SM estimates. Despite applying the same Change detection model, there are significant differences between the two products. These differences between the model using S-1 data and the model with ASCAT data are the different principle of vegetation correction due to the different methods of data collection of both sensors, the size of the dataset archive and the spatial resolution. Temporal and spatial correction of vegetation of the ASCAT model has already been studied and evaluated in several studies [Hahn et al., 2021b],[Pfeil et al., 2018], [Vreugdenhil et al., 2016]. S-1 correction is still a less researched topic and is the subject of future research. The length of the data archive can also cause complications with the S-1 model as stated by [Bauer-Marschallinger et al., 2019] due to possible failure to capture dry and wet soil conditions. The influence of spatial resolution can be demonstrated on the example of the *Cotnari* station. The ASCAT SM product shows a lower $RMSD$ (0.03 $m^3.m^{-3}$) for this location as well as, although still moderate, a higher correlation (ca. 45%) than the S-1 SM estimates. The polygon manually created to cover the nearby vineyard field next to *Cotnari* station covers the second smallest area (3 468 m^2). As described by [Bauer-Marschallinger et al., 2019] the small-scale scattering mechanisms can at field-scale contribute significantly to the backscatter coefficient and hence distort the time series. Upscaling the data to a coarser resolution can paradoxically help to improve the agreement between in-situ data and satellite estimates due to the suppression of small-scale effects.

When comparing the difference between all three products and the in-situ SM, it is shown that often both sensors (S-1 C-SAR and ASCAT C-band) fail to estimate measured SM at the same locations (Fig. 5.5). Even with the vegetation correction im-
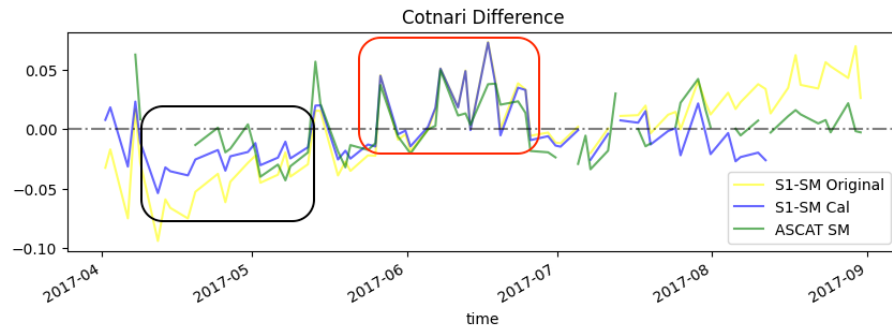
Figure 5.5 Comparison of SM estimates by original Change detection model, calibrated
model and ASCAT SM product

plemented in the ASCAT product, there is still a present overestimation during summer
months (red rectangle) as well as for the S-1 SM time series, however the magnitude of
the difference is mostly smaller and closer to 0 (dashed black horizontal line) indicating
a stronger impact of vegetation correction for the ASCAT model than implemented in
the current thesis. In the black rectangle the above mentioned underestimation during
spring months is highlighted showing that calibration even if not outperformed the
ASCAT SM it reaches comparable $RMSD$ values.

Conclusion

The presented thesis deals with the determination of soil moisture from S-1 SAR imagery at the individual field level using the SM Change detection model. The thesis proposes the calibration of this model in order to correct for the influence of vegetation in the temporal dimension, by creating a new dry reference based on CR to capture vegetation growth and development. The model calibration achieved an average improvement of -7.5% in $RMSD$ and almost 8.8% in correlation results. The results vary between stations however the calibration particularly benefited sites where the original model failed. The new dry reference in the model was able to reduce the negative and positive bias that results from the presence of vegetation. Despite the assumption that the largest difference between the model without and with vegetation correction would be observable during the season defined by the optical VPP data, the soil moisture estimates were improved especially during the off-season. This is likely due to the different SAR and VPP data acquisition processes.

The two models were also compared with the ASCAT SM product. Although the model calibration improved the results, the SM estimates using ASCAT are still more accurate. Nevertheless, the results of the thesis indicate the benefit of vegetation correction using CR in the SM Change detection model. In this work, the Change detection models are still burdened by the effect of insufficient LIA correction, which causes signal fluctuation between orbits. A possible solution would be to use data from only one relative orbit, but this would result in a decrease in the amount of data. The vegetation correction also shows different results for different years depending on the given conditions such as temperature and rainfall. The proposed correction is the same for all years and does not account for individual inter-annual specificities. The results of the work are based only on data in Romania and therefore may not be representative for larger than regional scales.

Local factors such as crop type, the influence of temperature and precipitation, surface roughness factor or soil composition as well as their influence on SM retrieval would need to be further analyzed in the future research. Additionally, implementation of the procedure could be interesting and beneficial for lower spatial resolution data which would not be burdened by small-scale scattering mechanisms. For the spatial analysis of the correction, it would be beneficial to use calibration data from a denser network spread over a smaller area and thus evaluate the results of the correction in

space.

Bibliography

[Attema and Ulaby, 1978] Attema, E. P. W. and Ulaby, F. T. (1978). Vegetation modeled as a water cloud. *Radio Science*, 13(2):357–364.

[Babaeian et al., 2019] Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., and Tuller, M. (2019). Ground, proximal, and satellite remote sensing of soil moisture. *Reviews of Geophysics*, 57(2):530–616.

[Baghdadi and Zribi, 2006] Baghdadi, N. and Zribi, M. (2006). Evaluation of radar backscatter models iem, oh and dubois using experimental observations. *International Journal of Remote Sensing*, 27:3831–3852.

[Balenzano et al., 2021] Balenzano, A., Mattia, F., Satalino, G., Lovergine, F., Palmisano, D., Peng, J., Marzahn, P., Wegmüller, U., Cartus, O., Dabrowska-Zielinska, K., Musial, J., Davidson, M., Pauwels, V., Cosh, M., McNairn, H., Johnson, J., Walker, J., Yueh, S., Entekhabi, D., and Balenzano, A. (2021). Sentinel-1 soil moisture at 1 km resolution: a validation study. *Remote Sensing of Environment*, 263:112554.

[Barrett et al., 2009] Barrett, B. W., Dwyer, E., and Whelan, P. (2009). Soil moisture retrieval from active spaceborne microwave observations: An evaluation of current techniques. *Remote Sensing*, 1(3):210–242.

[Bauer-Marschallinger et al., 2021] Bauer-Marschallinger, B., Cao, S., Navacchi, C., Freeman, V., Reuß, F., Geudtner, D., Rommen, B., Vega, F. C., Snoeij, P., Attema, E., Reimer, C., and Wagner, W. (2021). The normalised sentinel-1 global backscatter model, mapping earth's land surface with c-band microwaves. 8(1):277. Publisher: Nature Publishing Group.

[Bauer-Marschallinger et al., 2019] Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., and Wagner, W. (2019). Toward global soil moisture monitoring with sentinel-1: Harnessing assets and overcoming obstacles. 57(1):520–539. Conference Name: IEEE Transactions on Geoscience and Remote Sensing.

[Bhogapurapu et al., 2022] Bhogapurapu, N., Dey, S., Homayouni, S., Bhattacharya, A., and Rao, Y. (2022). Field-scale soil moisture estimation using sentinel-1 grd sar

data. *Advances in Space Research*, 70(12):3845–3858. Advances in Spaceborne SAR Remote Sensing for Characterization of Natural and Manmade Features - Part 2.

[Boisvert et al., 1997] Boisvert, J. B., Gwyn, Q. H. J., Chanzy, A., Major, D., Brisco, B., and Brown, R. (1997). Effect of surface soil moisture gradients on modelling radar backscattering from bare fields. *International Journal of Remote Sensing*, 18(1):153–170.

[Bourbigot, 2016] Bourbigot, M. (2016). Sentinel-1 product definition. (2).

[Bowers and Smith, 1972] Bowers, S. A. and Smith, S. J. (1972). Spectrophotometric determination of soil water content. *Soil Science Society of America Journal*, 36(6):978–980.

[Dobson and Ulaby, 1986] Dobson, M. C. and Ulaby, F. T. (1986). Active microwave soil moisture research. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24(1):23–36.

[Dobson et al., 1985] Dobson, M. C., Ulaby, F. T., Hallikainen, M. T., and El-rayes, M. A. (1985). Microwave dielectric behavior of wet soil-part ii: Dielectric mixing models. *IEEE Transactions on Geoscience and Remote Sensing*, GE-23(1):35–46.

[Dorigo et al., 2021] Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., Baldocchi, D., Bitelli, M., Blöschl, G., Bogena, H., Brocca, L., Calvet, J.-C., Camarero, J. J., Capello, G., Choi, M., Cosh, M. C., van de Giesen, N., Hajdu, I., Ikonen, J., Jensen, K. H., Kanniah, K. D., de Kat, I., Kirchengast, G., Kumar Rai, P., Kyrouac, J., Larson, K., Liu, S., Loew, A., Moghaddam, M., Martínez Fernández, J., Mattar Bader, C., Morbidelli, R., Musial, J. P., Osenga, E., Palecki, M. A., Pellarin, T., Petropoulos, G. P., Pfeil, I., Powers, J., Robock, A., Rüdiger, C., Rummel, U., Strobel, M., Su, Z., Sullivan, R., Tagesson, T., Varlagin, A., Vreugdenhil, M., Walker, J., Wen, J., Wenger, F., Wigneron, J. P., Woods, M., Yang, K., Zeng, Y., Zhang, X., Zreda, M., Dietrich, S., Gruber, A., van Oevelen, P., Wagner, W., Scipal, K., Drusch, M., and Sabia, R. (2021). The international soil moisture network: serving earth system science for over a decade. 25(11):5749–5804. Publisher: Copernicus GmbH.

[Dubois et al., 1995] Dubois, P., van Zyl, J., and Engman, T. (1995). Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing*, 33(4):915–926.

[FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009] FAO/IIASA/ISRIC/ISS-CAS/JRC (2009). Harmonized world soil database - version 1.1. Technical report, FAO, Rome, Italy and IIASA, Laxenburg, Austria.

[Gauthier et al., 1998] Gauthier, Y., Bernier, M., and Fortin, J.-P. (1998). Aspect and incidence angle sensitivity in ers-1 sar data. *International Journal of Remote Sensing*, 19(10):2001–2006.

[GEE, nda] GEE (n.d.a). Harmonized sentinel-2 MSI: MultiSpectral instrument, level-2a | earth engine data catalog. accessed: 2024-05-27.

[GEE, ndb] GEE (ndb). MODIS terra daily NDVI | earth engine data catalog | google for developers.

[GEE, ndc] GEE (n.d.c). Sentinel-1 algorithms | google earth engine.

[Gruber et al., 2020] Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.-C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., van der Schalie, R., Wigneron, J.-P., and Wagner, W. (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors? *Remote Sensing of Environment*, 244:111806.

[H SAF, 2022] H SAF (2022). *Product User Manual (PUM) Metop ASCAT Surface Soil Moisture Climate Data Record v7 12.5 km sampling (H119) and Extension (H120).* v0.2.

[Hahn et al., 2021a] Hahn, S., Paulik, C., Schmitzer, M., and iteubner (2021a). TUW-GEO/ascat: Add h SAF FTP and EUMETSAT data store download interface.

[Hahn et al., 2021b] Hahn, S., Wagner, W., Steele-Dunne, S. C., Vreugdenhil, M., and Melzer, T. (2021b). Improving ASCAT soil moisture retrievals with an enhanced spatially variable vegetation parameterization. 59(10):8241–8256. Conference Name: IEEE Transactions on Geoscience and Remote Sensing.

[Hallikainen et al., 1985] Hallikainen, M. T., Ulaby, F. T., Dobson, M. C., El-rayes, M. A., and Wu, L.-k. (1985). Microwave dielectric behavior of wet soil-part 1: Empirical models and experimental observations. *IEEE Transactions on Geoscience and Remote Sensing*, GE-23(1):25–34.

[Holah et al., 2005] Holah, N., Baghdadi, N., Zribi, M., Bruand, A., and King, C. (2005). Potential of asar/envisat for the characterization of soil surface parameters over bare agricultural fields. *Remote Sensing of Environment*, 96(1):78–86.

[Kahle, 1977] Kahle, A. B. (1977). A simple thermal model of the earth's surface for geologic mapping by remote sensing. *Journal of Geophysical Research (1896-1977)*, 82(11):1673–1680.

[Karthikeyan et al., 2017] Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. (2017). Four decades of microwave satellite soil moisture observations: Part 1. a review of retrieval algorithms. *Advances in Water Resources*, 109:106–120.

[Khabbazan et al., 2019] Khabbazan, S., Vermunt, P., Steele-Dunne, S., Ratering Arntz, L., Marinetti, C., van der Valk, D., Iannini, L., Molijn, R., Westerdijk, K., and van der Sande, C. (2019). Crop monitoring using sentinel-1 data: A case study from the netherlands. *Remote Sensing*, 11(16).

[Kornelsen and Coulibaly, 2013] Kornelsen, K. C. and Coulibaly, P. (2013). Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. *Journal of Hydrology*, 476:460–489.

[Kweon and Oh, 2015] Kweon, S.-K. and Oh, Y. (2015). A modified water-cloud model with leaf angle parameters for microwave backscattering from agricultural fields. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5):2802–2809.

[Li et al., 2021] Li, Z.-L., Leng, P., Zhou, C., Chen, K.-S., Zhou, F.-C., and Shang, G.-F. (2021). Soil moisture retrieval from remote sensing measurements: Current knowledge and directions for the future. *Earth-Science Reviews*, 218:103673.

[Liu and Yang, 2022] Liu, Y. and Yang, Y. (2022). Advances in the quality of global soil moisture products: A review. *Remote Sensing*, 14(15).

[Lv et al., 2018] Lv, S., Zeng, Y., Wen, J., Zhao, H., and Su, Z. (2018). Estimation of penetration depth from soil effective temperature in microwave radiometry. *Remote Sensing*, 10(4).

[Ma et al., 2020] Ma, C., Li, X., and McCabe, M. F. (2020). Retrieval of high-resolution soil moisture through combination of sentinel-1 and sentinel-2 data. *Remote Sensing*, 12(14).

[Ma et al., 2024] Ma, Y., Jiang, G., Huang, J., Shen, Y., Guan, H., Dong, Y., Li, J., and Hu, C. (2024). Evaluating the ability of the sentinel-1 cross-polarization ratio to detect spring maize phenology using adaptive dynamic threshold. *Remote Sensing*, 16(5).

[Mialon et al., 2015] Mialon, A., Richaume, P., Leroux, D., Bircher, S., Bitar, A. A., Pellarin, T., Wigneron, J.-P., and Kerr, Y. H. (2015). Comparison of dobson and mironov dielectric models in the smos soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 53(6):3084–3094.

[Mironov et al., 2009] Mironov, V. L., Kosolapova, L. G., and Fomin, S. V. (2009). Physically and mineralogically based spectroscopic dielectric model for moist soils. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2059–2070.

[Mo et al., 1982] Mo, T., Choudhury, B. J., Schmugge, T. J., Wang, J. R., and Jackson, T. J. (1982). A model for microwave emission from vegetation-covered fields. *Journal of Geophysical Research: Oceans*, 87(C13):11229–11237.

[Mullissa et al., 2021] Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., Gou, Y., Gorelick, N., and Reiche, J. (2021). Sentinel-1 sar backscatter analysis ready data preparation in google earth engine. *Remote Sensing*, 13(10).

[Naeimi et al., 2009] Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W. (2009). An improved soil moisture retrieval algorithm for ers and metop scatterometer observations. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):1999–2013.

[Njoku and Entekhabi, 1996] Njoku, E. G. and Entekhabi, D. (1996). Passive microwave remote sensing of soil moisture. *Journal of Hydrology*, 184(1):101–129. Soil Moisture Theories and Observations.

[Oh et al., 1992] Oh, Y., Sarabandi, K., and Ulaby, F. (1992). An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):370–381.

[Oki and Kanae, 2006] Oki, T. and Kanae, S. (2006). Global hydrological cycles and world water resources. *Science*, 313(5790):1068–1072.

[Paluba, 2020] Paluba, D. (2020). A correction of the local incidence angle of SAR data: a land cover specific approach for time series analysis. Publisher: Unpublished.

[Peel et al., 2007] Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the köppen-geiger climate classification. 11(5):1633–1644. Publisher: Copernicus GmbH.

[Petropoulos et al., 2015] Petropoulos, G. P., Ireland, G., and Barrett, B. (2015). Surface soil moisture retrievals from remote sensing: Current status, products and future trends. *Physics and Chemistry of the Earth, Parts A/B/C*, 83-84:36–56. Emerging science and applications with microwave remote sensing data.

[Pfeil et al., 2018] Pfeil, I., Vreugdenhil, M., Hahn, S., Wagner, W., Strauss, P., and Blöschl, G. (2018). Improving the seasonal representation of ascat soil moisture and vegetation dynamics in a temperate climate. *Remote Sensing*, 10(11).

[Preimesberger et al., 2023] Preimesberger, W., Pbutting, Daberer, Tracyscanlon, Schmitzer, M., D. Baum, Lzappa, Hahn, S., Bascrezee, Paulik, C., and Bader, N. F. (2023). TUW-GEO/ismn: v1.3.4.

[Sadeghi et al., 2017] Sadeghi, M., Babaeian, E., Tuller, M., and Jones, S. B. (2017). The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to sentinel-2 and landsat-8 observations. *Remote Sensing of Environment*, 198:52–68.

[Seneviratne et al., 2010] Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3):125–161.

[Sentinel-1, nd] Sentinel-1 (n.d.). User guides - sentinel-1 SAR - overview - sentinel online.

[Smets et al., 2023a] Smets, B., Cai, Z., Eklundh, L., Tian, F., Bonte, K., Van Hoost, R., De Roo, B., Jacobs, T., Camacho, F., Sánchez-Zapero, J., Martinez-Sánchez, E., Swinnen, E., Scheifinger, H., Hufkens, K., and Jönsson, P. (2023a). Hr-vpp product user manual seasonal trajectories and vpp parameters.

[Smets et al., 2023b] Smets, B., De Roo, B., Jacobs, T., and Daems, D. (2023b). High resolution vegetation phenology and productivity (hr-vpp) data access.

[Trenberth et al., 2009] Trenberth, K. E., Fasullo, J. T., and Kiehl, J. (2009). Earth's global energy budget. *Bulletin of the American Meteorological Society*, 90(3):311 – 324.

[Ulaby et al., 1978] Ulaby, F. T., Batlivala, P. P., and Dobson, M. C. (1978). Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: Part i-bare soil. *IEEE Transactions on Geoscience Electronics*, 16(4):286–295.

[Ulaby et al., 1990] Ulaby, F. T., Sarabandi, K., McDonald, K., Whitt, M., and Dobson, M. C. (1990). Michigan microwave canopy scattering model. *International Journal of Remote Sensing*, 11(7):1223–1253.

[van Oevelen and Hoekman, 1999] van Oevelen, P. and Hoekman, D. (1999). Radar backscatter inversion techniques for estimation of surface soil moisture: Efeda-spain and hapex-sahel case studies. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):113–123.

[Veloso et al., 2017] Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., and 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.

[Verhoest et al., 2008] Verhoest, N. E., Lievens, H., Wagner, W., Álvarez Mozos, J., Moran, M. S., and Mattia, F. (2008). On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. *Sensors*, 8(7):4213–4248.

[Vreugdenhil et al., 2016] Vreugdenhil, M., Dorigo, W. A., Wagner, W., de Jeu, R. A. M., Hahn, S., and van Marle, M. J. E. (2016). Analyzing the vegetation parameterization in the tu-wien ascat soil moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6):3513–3531.

[Vreugdenhil et al., 2020] Vreugdenhil, M., Navacchi, C., Bauer-Marschallinger, B., Hahn, S., Steele-Dunne, S., Pfeil, I., Dorigo, W., and Wagner, W. (2020). Sentinel-1 cross ratio and vegetation optical depth: A comparison over europe. 12(20):3404. Number: 20 Publisher: Multidisciplinary Digital Publishing Institute.

[Wagner et al., 2019] Wagner, W., Gruber, A., Hahn, S., Vreugdenhil, M., Melzer, T., and Quast, R. (2019). Data retrieval in earth observation. page 72.

[Wagner et al., 2013] Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-saldaña, J., Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., and Rubel, F. (2013). The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. 22:5–33.

[Wagner et al., 1999] Wagner, W., Lemoine, G., and Rott, H. (1999). A method for estimating soil moisture from ers scatterometer and soil data. *Remote Sensing of Environment*, 70(2):191–207.

[Wang and Schmugge, 1980] Wang, J. R. and Schmugge, T. J. (1980). An empirical model for the complex dielectric permittivity of soils as a function of water content. *IEEE Transactions on Geoscience and Remote Sensing*, GE-18(4):288–295.

[Yu et al., 2023] Yu, H., Yang, Y., Wang, C., Chen, R., Xie, Q., Liu, G., and Yin, G. (2023). Extracting deciduous forests spring phenology from sentinel-1 cross ratio index. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16:2841–2850.

[Zhang and Zhou, 2016] Zhang, D. and Zhou, G. (2016). Estimation of soil moisture from optical and thermal remote sensing: A review. *Sensors (Basel, Switzerland)*, 16.

[Zhu et al., 2022] Zhu, L., Si, R., Shen, X., and Walker, J. (2022). An advanced change detection method for time-series soil moisture retrieval from sentinel-1. *Remote Sensing of Environment*, 279:113137.

[Ångström, 1925] Ångström, A. (1925). The albedo of various surfaces of ground. *Geografiska Annaler*, 7:323–342.